APPLICATION OF LASER RAMAN SPECTROSCOPY IN CONFORMATIONAL ANALYSIS OF PROTEINS

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

by

Byeong Hyeok Jo

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy in the

School of Chemistry

Georgia Institute of Technology

July 1975

APPLICATION OF LASER RAMAN SPECTROSCOPY

IN CONFORMATIONAL ANALYSIS OF PROTEINS

Approved:

Nai-Teng Yu, Chairman

Sidney L. Gordon

Raymond F. Borkman

Date approved by Chairman: July 11 '75

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to my thesis advisor, Professor Nai-Teng Yu, for his understanding, advice and endless patience during the course of this research. A special note of thanks is due Professor D. C. O'Shea for granting me the use of his Raman Spectrometer and his assistance in instrument operation. I would like to thank Professors S. L. Gordon and R. F. Borkman for valuable suggestions in their review of this manuscript and for serving as members of my reading committee. I also wish to thank Dr. E. J. East for reading and correcting the English of an initial draft of the manuscript.

I wish to express my sincere appreciation to my parents whose love and guidance will always be cherished.

I am greatly indebted to my wife, Soon-Ae, for her sacrificial and untiring support and encouragement during the course of this work. Words alone are so inadequate to express my thanks.

I am also grateful to the Korea Society, Inc. which awarded me DeWitt Wallace Fellowship and the NIH which provided me a research assistantship.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

CHAPTER III

CHAPTER IV

LIST OF ILLUSTRATIONS

CHAPTER I

LIST OF ILLUSTRATIONS (Continued)

CHAPTER II

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

SUMMARY

Laser Raman spectroscopy was applied to analyzing protein conformation in a variety of sample phases. As model compounds, lysozyme and ribonuclease A(RNase A) were selected. The spectra of the crystalline state were interpreted on the basis of X-ray models and compared to those obtained in solution, powder and denatured state.

Some of the findings are that the lysozyme molecule has the same backbone conformation in aqueous solution as in the crystalline state and the same is true for RNase A, but some of the side-chains in both proteins have subtle differences in conformation between the two phases.

The Raman spectral regions on which structural variations of proteins are sensitively monitored include S-S and C-S stretching vibrations, C-C-N skeletal stretching vibrations, amide I and amide III and several aromatic-ring vibrations of side-chains.

The knowledge obtained from the studies of the model compounds was used to analyze snake venom proteins such as cobramine B, neurotoxin α, Pelamis platurus major toxin and Laticauda semifasciata toxin b, carboxymethylated chains of insulin, and bovine C-peptide in terms of main-chain conformation and side-chain conformation. Raman spectra of carboxypeptidase A(CPDase A) and single crystal insulin were

presented and correlated to their X-ray structures. To complement the interpretation of the amide I and III regions of the spectra of snake venoms and bovine C-peptide, Chou and Fasman's conformation calculation was performed on neurotoxin α and bovine and porcine C-peptides. It was concluded that snake venoms retain a significant portion of random-coil relative to other components in the backbone conformation and C-peptides consist of random-coil and α-helices. Furthermore, the effect of urea on the lysozyme structure was reported on the basis of Raman spectra. No changes in the amide III region were observed at concentrations up to 6 M, but 9 M urea, small changes were observed. Small changes in the 600-750 cm^{-1} region were detected in the presence of 4 or 6 M urea.

CHAPTER I

COMPARISON OF LYSOZYME STRUCTURES BETWEEN SOLUTION AND CRYSTALLINE, LYOPHILIZED AND TRICHLOROACETIC ACID(TCA)-TREATED POWDERS

Introduction

The laser Raman-scattering technique was employed to detect the structure of lysozyme in various states such as crystals, solution, lyophilized powder and trichloroacetic acid(TCA)-treated solid. The lysozyme, an enzyme from egg white that catalyzes the hydrolysis of certain glycosidic linkages in the complex polysaccharides found in the cell walls of some bacteria, is a globular protein of molecular weight 14,600, with 129 amino acid residues in a single polypeptide chain and four disulfide bridges.

It is generally believed that the function of a globular protein in vivo depends on its detailed three-dimensional structure in aqueous solution. The determination of such structure in solution is of biological importance. However, because of the absence of techniques for doing so, X-ray diffraction studies have been carried out with protein crystals and it has been assumed that the protein structures are the same in crystals as in solution. The question of whether they are really the same has been

the subject of considerable investigation (Rupley 1969).

The assumption might be justified if one could consider protein crystals as similar to highly concentrated solutions since each protein molecule in the crystal is surrounded by solvent molecules except for a small number of protein-protein contacts (Edsall 1968). For example, there are six such contacts for hemoglobin (Perutz 1965) and five for lysozyme (Blake et al. 1967). When the protein molecules leave the crystal and go into the solution, the water layer around the molecular surface could presumably remain essentially unchanged, and the only change is the replacement of protein-protein contacts by protein-solvent contacts (Edsall 1968). Since the area of contacts between neighboring molecules is small as compared to the total surface area, the effect of dissolution on protein conformation might be negligible. However, one may argue that since the protein molecules are flexible (Nakanishi et al. 1972, Linderstrom-Land and Schellman 1959), it is not unlikely that the lattice force, although weak, may distort the protein conformation upon crystallization (Rupley 1969). Nevertheless, experiments that compare crystals and solution are necessary because of the possibility of small changes in conformation.

Raman studies have also been made on lyophilized and TCA-treated powders and TCA-treated solution. The lyophilization is a common process of reducing protein solution to

dryness. Through this process, the unbonded and loosely bonded water molecules are removed from the frozen sample with a concomitant replacement of a certain number of protein-solvent contacts by protein-protein contacts. TCA, a commonly used protein precipitant in analytical biochemistry, was suggested to denature proteins before it precipitates proteins (Borsook 1950, Takenaka et al. 1971). The interaction of TCA with a protein may involve a water extrusion mechanism which may occur in the interaction between COO^- group of TCA with lysine and arginine side-chains of the protein and hydrophobic interaction of CCl_3 group of TCA with hydrophobic side-chains of the protein (Lewin 1970). The structural alteration in protein associated with TCA denaturation could be followed conveniently by Raman spectroscopy. The solution of TCA-treated lysozyme was prepared to check whether the TCA denaturation of proteins is reversible or irreversible.

Recently Yu et al. (1972) made a preliminary comparison between the Raman spectra of crystalline and aqueous insulin and noted that in the so-called amide III region (1220-1300 cm^{-1}) there existed a small difference in line shape. Since the Raman amide III vibrations are known to be sensitive to the conformation of protein backbone (Yu and Liu 1972), the spectra then reflect a difference in backbone conformation between the two states. Since the spectra presented earlier contain considerable noise,

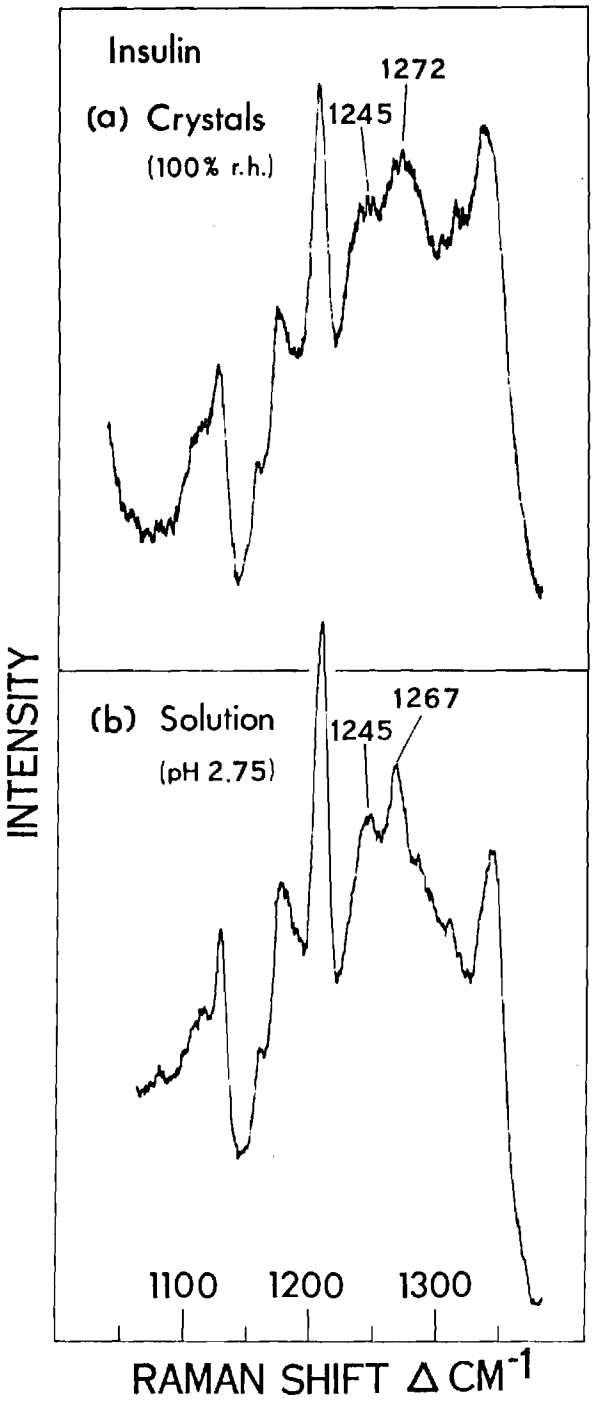

Figure 1. Raman Spectra of Insulin in Crystalline Powder and Solution States

improved spectra have been obtained and presented in Figure 1.

Materials and Methods

Uniform rhombohedral beef zinc-insulin crystals (about 0.025 mm) were obtained from Lilly's "Ultralente" U-80 insulin zinc suspension injection. To remove 0.1% methylparaben preservative, crystals were washed with distilled water three times by centrifugation and decantation of the supernatant solution. The wet crystals of insulin were then kept in a 100% relative humidity environment for about 10 hours before the experiments. The shape of the crystals was examined by microscope and found unchanged after long exposure (about 12 hours) to the laser beam. The insulin solution was prepared by dissolving 100 mg of dried insulin crystals in 1 ml of solution. Because of the low solubility of insulin at pH 7, the pH value of the solution was adjusted with 1.0 N HCl to 2.75. The solution was then centrifuged at a speed of 8000 rpm for 0.5 hours before injection to the capillary cells. No additional salts were added.

Crystals of hen egg-white lysozyme chloride were grown from 5% sodium chloride solution (about 0.85 M) at pH 4.5 according to the method of Alderton and Febold (1946). This was also the method used by Blake and co-workers (1967) in preparing the crystals for their X-ray diffraction

studies. The crystals obtained in our laboratory were of reasonably uniform sizes and shapes. The maximum dimension of crystals was about 0.075 mm. After being removed from the solvent, crystals were gently ground into smaller crystals (less than 0.005 mm) to ensure that the crystals were much smaller than the diameter of the focused laser beam (about 0.050 mm). Thus, the scattered light detected came from many randomly oriented crystals, instead of a single crystal. This will make the comparison of crystals and solution spectra more meaningful and avoid the possible dependence of scattering intensities on the orientation of the single crystal. The crystalline powder was then packed into a conical depression at the end of an 1/8-in. steel rod, which was then horizontally fastened in a Thermovac flask equipped with a rubber "O" ring and a vacuum-tight stopcock. The sample in the flask was equilibrated at room temperature with saturated H_2O vapor for 7 hours. The solution samples for experiments were prepared by dissolving 125 mg lysozyme chloride crystals in 0.5 ml solution. The pH values of the solutions were adjusted with 0.1 N HCl (or 0.1 N NaOH) and determined by a Radiometer pH meter model 26 with a microprobe combination electrode (Fisher Cat. No. 13-639-92). Since there are about 16 chloride ions per lysozyme molecule in the crystal, the chloride ions introduced into the solution were about 0.28 M. Of course, there was also a small amount of NaCl contaminated on the

surface of crystals. However, the ionic strength of the solution was much lower than that of crystal-growing medium.

The lyophilized powder of lysozyme was obtained by freeze-drying the above solution. It was packed into the depression of a steel rod just like that for crystalline powder. The sample was kept in the flask at 0% relative humidity(r.h.) for 2 days. The 0% r.h. was achieved by using phosphorous pentoxide and dry nitrogen.

The TCA-treated lysozyme powder was prepared by adding 5% TCA solution to the same solution used above. The precipitated lysozyme-TCA complex was then washed three times with 5 ml of acetone containing 0.1% hydrochloric acid. In this process, it is known that the trichloroacetate group in the complex is replaced by chloride ion (Hnilica et al. 1962). This is the method used in the process of histone fractionation (John and Butler 1972, Hnilica et al. 1962). It was air dried and therefrom the Raman spectrum was obtained. The spectrum is completely free from Raman bands of TCA. The solution of TCA-treated lysozyme was prepared by dissolving the TCA-treated lysozyme powder in water with conditions adjusted to those of native lysozyme solution used above. The spectrum of TCA-treated powder was taken in the same way as that of crystalline sample except that there was no 100% relative humidity.

The Raman spectrometer used for obtaining the spectra was the same as that described by Yu et al. (1972).

The laser beam was generated by a Coherent Radiation model 52B argon-ion laser. After passing through an interference filter which removed the plasma emission lines, the monochromatic beam at 514.5 nm wavelength was focused and directed upward onto the sample at the grazing angle so that the scattering column was a strip on the powder surface, 1/8-in. long and approximately 40 microns wide. The scattered light is collected by an f/1.1 lens and imaged with a 3:1 magnification on the entrance slit of a Spex 1401 double monochromator. The light is detected by an ITT "Startracker" (FW-k30) photomultiplier which is thermoelectrically cooled to reduce the dark count rate to a few counts per second. The photomultiplier signal is amplified and processed by standard nuclear counting electronics with suppression of a large amount of background signal. An analog signal of the count rate is recorded on a strip chart recorder at a rate of 10 cm^{-1}/min (or 5 cm^{-1}/min) with a time constant which gives a 1% deviation for full scale signal. Wave number accuracy is ± 2 cm^{-1} except for the broad bands like the amide III where it is ± 4 cm^{-1}. The laser power at the sample was kept as low as possible so long as the signal-to-noise ratio of the spectra was acceptable. This was about 150 to 300 mW.

For liquid samples, a Cary Instrument's kinematic base and backing mirror assembly (Cat. No. 8240150) was used to hold a capillary cell, which is 1 mm (internal diameter)

x 25 mm long and made of Pyrex. The cell with a fire-polished flat end contained about 10 μl of solution and was held vertically. The light scattered at 90° to the incident beam was then collected and focused onto the entrance slit of the double monochromater. Whenever there was a question of grating ghosts, the 488.0 nm line was used to repeat the region of the spectra under investigation since the location of grating ghosts generally depends of the frequency of the exciting line. The grating ghosts are indicated by the X mark in Figures 2 and 3. The degree of uncertainty in the intensity reproducibility of each spectrum is negligible when it is compared to the changes in relative intensity arising from conformational changes. To check whether the laser beam has any effect on the sample during obtaining spectrum, the scanning of each spectrum was repeated at least three times. No changes have been noticed. It was also reported that the lysozyme retained its entire activity after prolonged irradiation by the laser beam at high power (450 mW) up to 6 hours (Brunner and Sussner 1972).

Results and Discussion

In Figure 2, the Raman spectra of crystals and solution of lysozyme are compared. The amide III line shape in the crystal spectrum agrees nicely with that of the solution. The three resolved peaks at 1238, 1258 and 1272 cm^{-1} were identified earlier by Lord and Yu (1970a), but assignments

to specific structural features were not made, although they did suggest that these resolved peaks might be associated with α-helix, β-structure, and random-coil in lysozyme molecule. Recently, Yu and Liu (1972), based on their Raman studies of glucagon in various conformational states, concluded that the α-helical, random-coiled (H-bonded) and antiparallel-β structure of proteins should have the amide III vibrations at 1266, 1248 and 1232 cm^{-1} respectively. The X-ray diffraction studies with lysozyme crystals (Blake et al. 1967, Phillips 1966) reveal that the molecule contains three regions of α-helix—residues 5-15, 24-34 and 88-96, antiparallel β-pleated sheet—residues 41-45 and 50-54 and random-coil structure. On this basis, the author may assign the 1238, 1258 and 1272 cm^{-1} peaks to antiparallel-β, random-coil and α-helix, respectively. Because of the agreement between the two spectra in this region, the author concludes that the conformation of lysozyme backbone is the same in the crystalline powder as it is in solution. As to the amide I region, there is only a single peak at 1660 cm^{-1}, indicating that the amide I region is less sensitive in detecting the various structural components in protein backbone. Nevertheless, the author also observed good agreement in this region between the two phases.

When lysozyme was freeze-dried from the solution, the relative peak intensity of 1238 and 1272 cm^{-1} in the amide III region appear to be somewhat different from that of

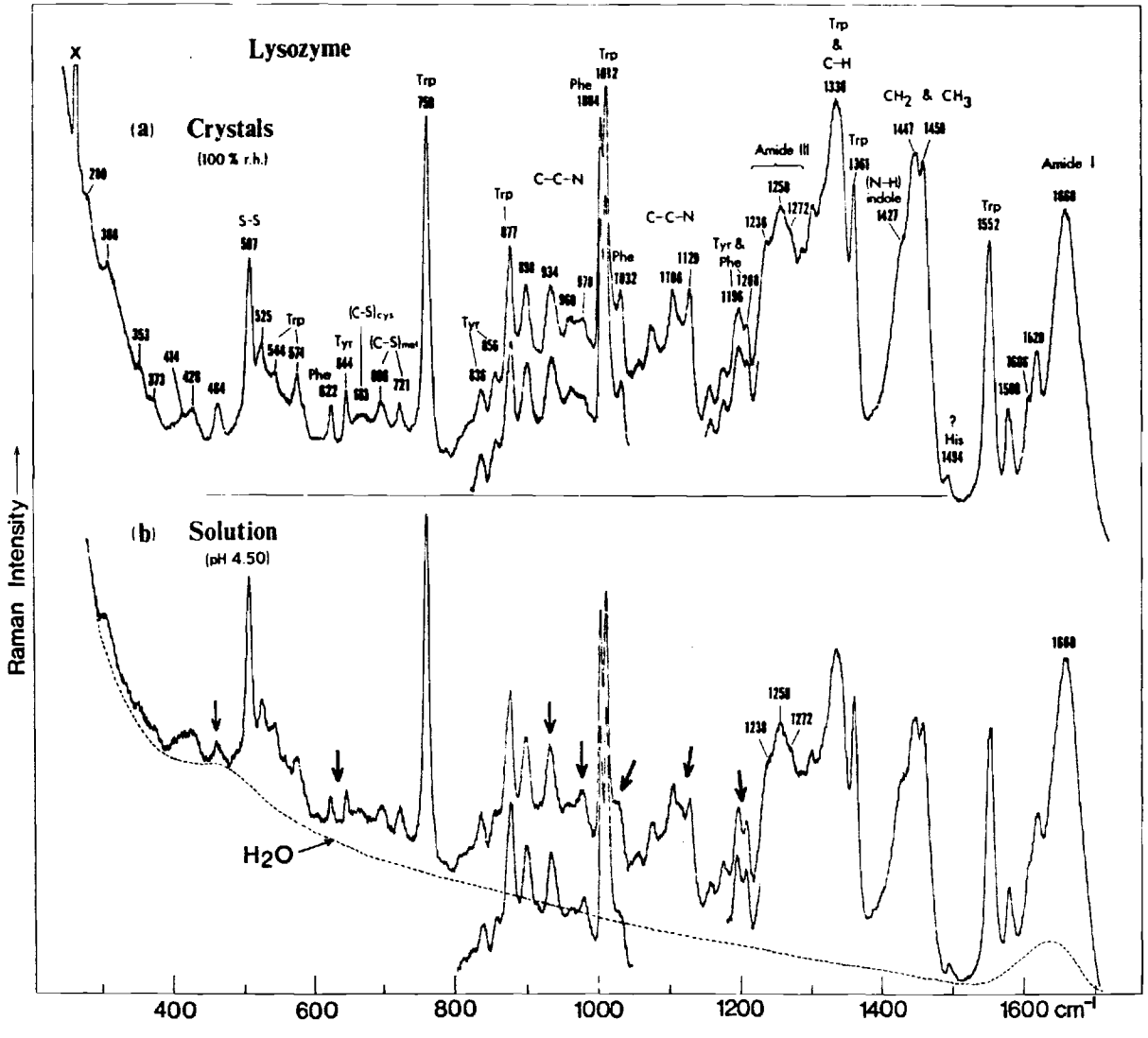

Figure 2. Raman Spectra of Lysozyme in Crystalline Powder and Solution States

crystalline powder and solution, indicating changes in main-chain conformation upon lyophilization. A Raman spectrum of TCA-denatured lysozyme powder is presented in Figure 4. In the amide III region, the band becomes sharpened and relatively symmetrical with the center of gravity at 1236 cm^{-1}. The sharpening of the amide III may indicate the uniformity of H-bonds. The amide I line has shifted from 1660 cm^{-1} to 1672 cm^{-1} and has sharpened considerably upon TCA denaturation. Similar spectral changes have also been observed by Chen et al. (1973, 1974) and Yu et al. (1972) upon chemical denaturation of lysozyme and the heat denaturation of insulin. Yu et al. (1972) suggested that as one factor the coupling between the adjacent peptide units (both intra- and inter-chain coupling) should involve in determining the amide I vibration of the Raman as well as that of the infrared and explained the shift to 1672 cm^{-1} as a rise due to one coupled mode. From the results of Raman studies of model compounds of polyglycine I (Small et al. 1970), poly-L-lysine (Wallách et al. 1970) and glucagon (Yu and Liu 1972), Yu et al. (1972) assigned the amide I line, 1672 cm^{-1}, to antiparallel-β structure. Contrary to this assignment, Chen et al. (1973, 1974) assigned this sharp amide I band to random-coil conformation since the amide III region in the denatured lysozyme spectra have a peak at a relatively higher frequency than that of model compound spectra of β structure. In the spectrum of

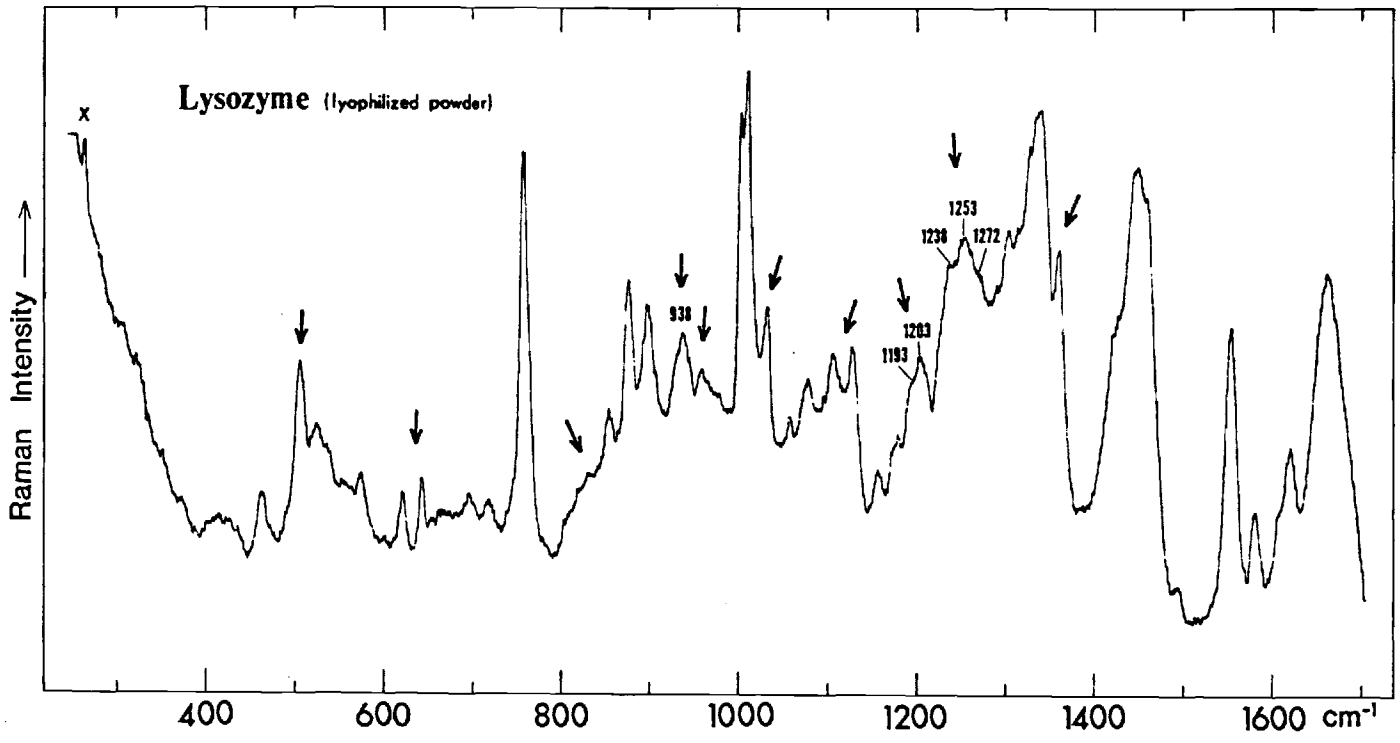

Figure 3. Raman Spectrum of Lysozyme in Lyophilized Powder

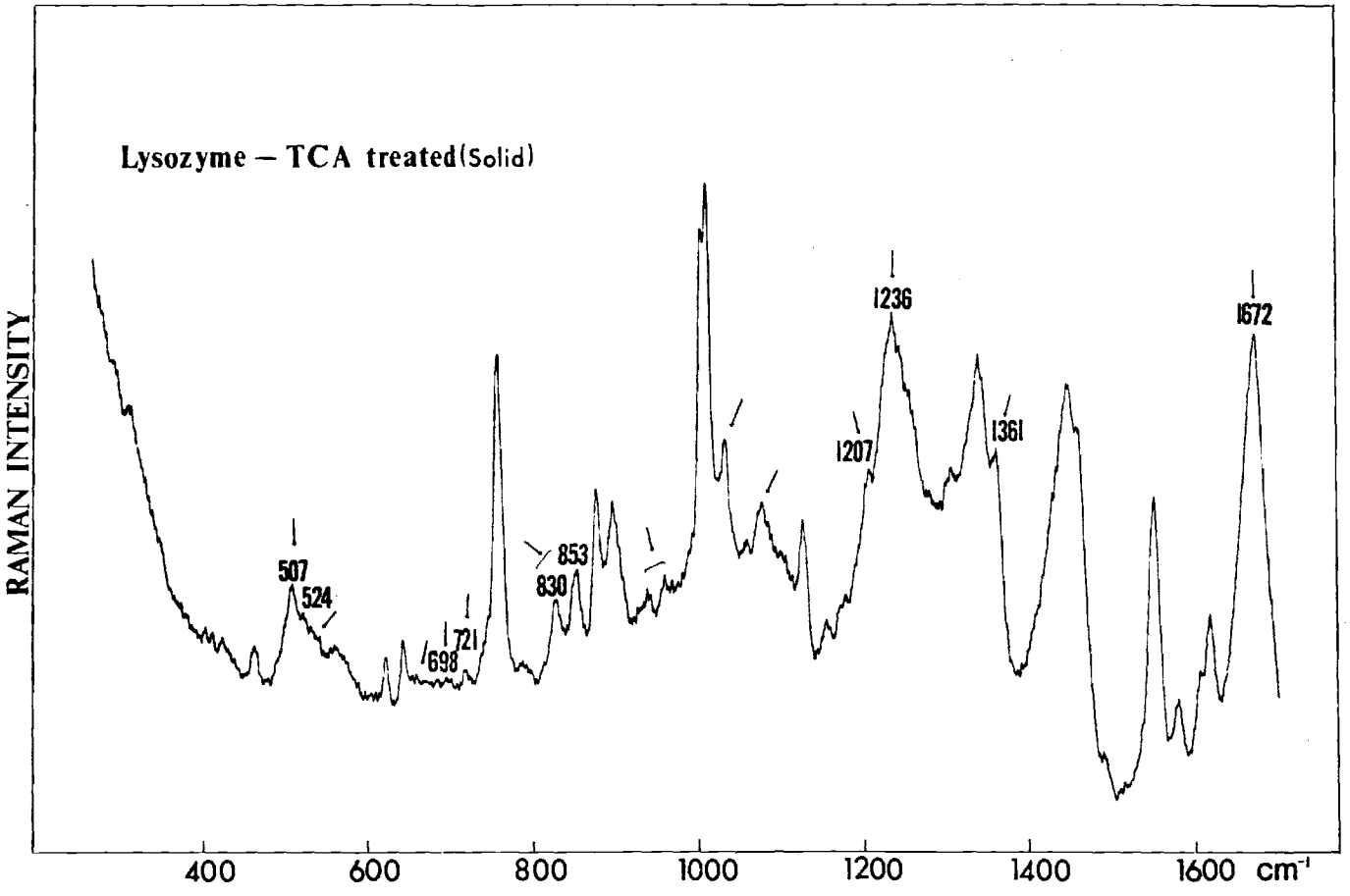

Figure 4. Raman Spectrum of TCA-treated Lysozyme Powder

TCA-denatured lysozyme, the amide III peak (1236 cm^{-1}) has not shifted much from the 1232 cm^{-1} band at amide III assigned to the β structure components of native lysozyme. Therefore, the author tentatively concludes that TCA-denatured lysozyme in the solid state may exist in predominantly antiparallel-β backbone conformation, although further studies should be pursued on this problem.

In the spectrum of Figure 2, noteworthy spectral changes occurred near 464, 622, 644, 934, 960, 978, 1032, 1129 and 1196 cm^{-1}, as indicated by arrows. It is rather unusual that several Raman lines (eg. 934, 1004, 1012 and 1196 cm^{-1}) are sharper in the aqueous phase than in the crystalline phase. Normally the opposite effect is observed with amino acids and simple peptides (Yu 1969). However, the line at 464 cm^{-1} broadened upon dissolution. On the other hand, two phenylalanyl lines at 622 and 1032 cm^{-1} and one tyrosyl line at 644 cm^{-1} weakened in their intensities when lysozyme dissolved.

When lysozyme was freeze-dried from the solution, numerous spectral changes were observed (Figure 3). The S-S line at 507 cm^{-1} broadened and the lines at 622 and 644 cm^{-1} increased their intensities. The lines at 759, 1004 and 1012 cm^{-1} were broader in the spectrum of the lyophilized powder than that of solution. The one at 934 cm^{-1} was shifted to 938 and broadened, and the doublet at 960 and 978 cm^{-1} changed the relative intensity. The 1032 cm^{-1}

line due to phenylalanyl rings increased its intensity. Other striking changes occurred near 835 and 1193 cm^{-1}.

Upon TCA denaturation, the line at 507 cm^{-1} decreased and became broader than that of lyophilized powder. The intensity drop at 525, 544, 574, 663, 696 and 721 cm^{-1} are also marked. The tyrosyl lines at 830 and 853 cm^{-1} became clearly defined. The skeletal frequencies involving the C-C and C-N stretching vibration show drastic changes. The drop at 1361 cm^{-1} is noticeable, too.

According to the X-ray diffraction studies of lysozyme crystals (Blake et al. 1967), there are five areas of surface contact between neighboring molecules. The first intermolecular interaction involves a double salt bridge between the negatively charged terminal carboxyl group and positively charged Lys 13 group of one molecule with the corresponding groups on the adjacent molecule. The second interaction is that the guanidino group of Arg 114 is sandwiched between Phe 34 of the same molecule and Tyr 23 of the next. The remaining interactions involve quite a few hydrogen bonds: between the main chain CO group of 113 and the γ-amino group of Asn 106 of the next molecule; between the main chain CO of 22 and the $N\eta$ of Arg 114 of the neighboring molecule; between the γ-amino group of Asn 39 and the main chain carbonyl groups of residues 66 and 67 of the next molecule; from the δ-amino group of Gln 41 to the main chain CO group of Gln 41 from the main chain NH group of residue

80 of the next molecules. The replacement of these hydrogen bonds by protein-solvent hydrogen bond is expected to broaden those vibrations directly derived from surface residues involved in protein-protein contacts. The broadening of the line at 464 cm^{-1} might be due to this fact. The 464 cm^{-1} line in Figure 3 for lyophilized powder became sharper than that for crystals. This may be due to the fact that freeze-drying process removes the water molecules existing around the protein-protein contacting areas. The 464 cm^{-1} line in Figure 4 is sharp but decreased in peak intensity. The TCA interaction with proteins was assumed to extrude water molecules binding to lysine and arginine side chains in addition to hydrophobic effect of trichloro group (Lewin 1970). Therefore, the unfolding of backbone conformation to antiparallel-β-pleated structure and the water molecule extrusion around lysine and arginine groups in the protein-protein contacting area are thought to result in a new shape of Raman line at 464 cm^{-1}. Although the exact nature of this line is not known, the author may tentatively assign it to the bending vibrations of the surface group which is sensitive to hydration effect.

Since Phe 34 is involved in intermolecular interaction, the intensity of phenylalanyl lines may be affected when protein-protein contact at this site is replaced by water solvation. In order to determine whether this is true,

the author has compared Raman spectra of glycyl-L-phenylalanyl-L-alanine both in the solid state and in solution. The intensities of the lines at 622 and 1032 cm^{-1} relative to the one at 1004 cm^{-1} did decrease upon dissolution. Thus, the decrease of the intensities at 622 and 1032 cm^{-1} due to phenyl rings in the spectra of lysozyme may be due to this solvation effect and is not necessarily associated with the conformational change. The relative intensities of 622 and 1032 cm^{-1} to 1004 cm^{-1} in the spectrum of lyophilized powder are not noticeably different from that for TCA-treated lysozyme. This supports the above speculation concerning the 1032 cm^{-1} line. Chen et al. (1974) also observed the increase in 1030 cm^{-1} in chemical denaturation of LiCl. It may be due to reorganization of water structure around the phenylalanyl group caused by high concentration of lithium ions.

The intensity drop at 644 cm^{-1} upon dissolution may be related to the same phenomenon observed in the spectra of lyophilized powder and aqueous RNase A (Yu et al. 1972). In the spectra of RNase A, Yu et al. (1972) found that the intensities of the tyrosine lines at 644 and 852 cm^{-1} relative to the one at 832 cm^{-1} decreased on dissolution. Because the direction of the intensity changes was opposite to that expected for solvation of tyrosyl rings, they suggested that this might be due to the changes in the local environment of the three "buried" tyrosines in RNase A.

Recently, Yu et al. (1973) studied the intensities of the tyrosyl lines in the spectra of cobramine B, a basic protein from cobra venom, and found that the intensity at 644 cm^{-1} also decreased upon dissolution. Since there is an evidence that all three tyrosines in cobramine B may be buried, it seems clear that the intensity drops at 644 and 856 cm^{-1} may be related to side-chain conformational changes. Through the freeze-drying process, the side-chain conformation is expected to change because of partial dehydration and protein-protein contact interaction. The line at 835 cm^{-1} in Figure 3 shows a striking decrease in intensity. The side-chain conformation of the tyrosyl groups in lysozyme seems to be affected little by TCA treatment as shown in the spectrum of Figure 4.

Concerning the tryptophan line at 1361 cm^{-1}, it was discussed that the indole rings accessible to water do not contribute to the distinct sharp feature on the basis of the Raman spectra of random-coiled glucagon (Yu and Liu 1972) and amino acid tryptophan in solution (Chen et al. 1973). In lysozyme, 3 to 4 tryptophan residues may be able to come in contact with the solvent (Hamaguchi and Kurono 1963, Takenaka et al. 1971). In other words, two or three buried tryptophan might contribute to the line of 1361 cm^{-1} in solution. From the degree of decrease in intensity of that line in the spectrum of lyophilized powder, it may be estimated that one more tryptophan residue became exposed

while TCA denaturation made about two tryptophans accessible to surface. In case of α-lactalbumin (Yu 1974), the 1361 cm^{-1} line shows a sharp feature and it was reported as a result of two buried tryptophan residues. Chen et al. (1974) observed decrease in intensity of 1361 cm^{-1} line in LiBr denaturation and assumed it as due to possible H-bond formation between the indole NH and bromide ion. In the author's opinion, the newly formed H-bond might have an effect on the position of a buried tryptophan by reorganizing the structure of solvent surrounding that residue. The decrease in that line also was observed in thermal denaturation by Brunner and Sussner (1972) to the same degree as in TCA or LiBr denaturation.

Other valuable indications of conformational variation are changes in the region of S-S and C-S stretching vibrations. The characteristics of the S-S and C-S stretching Raman bands were extensively studied using model compounds (Sugeta et al. 1972, 1973, Miyazawa and Sugeta 1974, Nogami et al. 1975). The origin of the S-S line near 510 cm^{-1} was assigned to the gauch-gauch-gauch form about the disulfide bond, while the line at 525 and 540 cm^{-1} to the gauch-gauch-trans and trans-gauch-trans forms respectively. A schematic diagram of geometry about the disulfide bond in the three forms is presented in Figure 5.

The molecular conformation of lysozyme about the disulfide bond in crystal was correlated with these lines

by Miyazawa and Sugeta (1974); three crystine groups are in the gauch-gauch-gauch form (64-80, 30-115, 6-127) and one in the gauch-gauch-trans form (94-76). Thus the line at 544 cm^{-1} in Figure 2(a) is thought to be contributed only by tryptophan residue. Little change is observed in this range between the crystalline powder and solution phases. On lyophilization, the peak of 525 cm^{-1} has become broadened to the top of the 544 cm^{-1} tryptophan peak, while the line at 507 cm^{-1} shows decrease in intensity. This suggests that a number of the gauch-gauch-gauch form might transform into a new type whose conformation lies between the gauch-gauch-trnas and trans-gauch-trans forms. After TCA treatment, there occurrs a drastic decrease and broadening in 507 cm^{-1} intensity making those lines at 524, 544 and 574 cm^{-1} less defined. This may be the result of several different allowed conformations of the four disulfide group rather than breaking some of disulfide bonds.

The C-S stretching vibration of cystine (Sugeta et al. 1972, 1973) and methionine (Nogami et al. 1975) groups were discussed using model compounds and Miyazawa and Sugeta (1974) made correlation with lysozyme structure. Their finding is that C-S stretching frequencies of primary disulfides $X\text{-}\overset{\prime}{\underset{|}{C}}\text{-}CH_2\text{-}S\text{-}S$ depend upon the atom (X) at the trans site with respect to the sulfur atom about the C-C bond. When X is a hydrogen atom (P_H), the C-S stretching vibration lies at 630-670 cm^{-1}, at 720 cm^{-1} for the conformation with

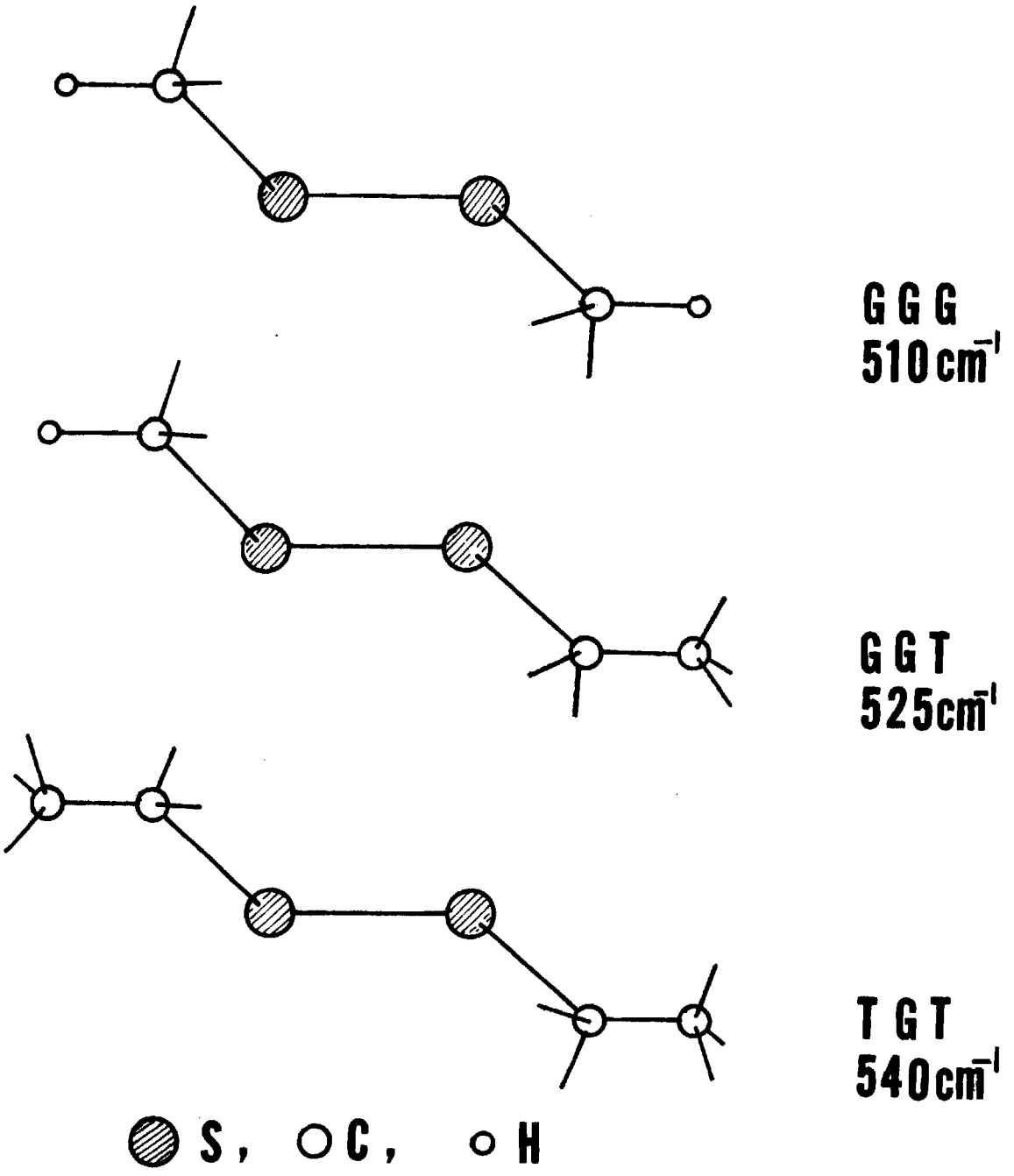

Figure 5. Correlation of Geometry about S-S Bond and Raman S-S Stretching Vibration (after Sugeta et al. 1972)

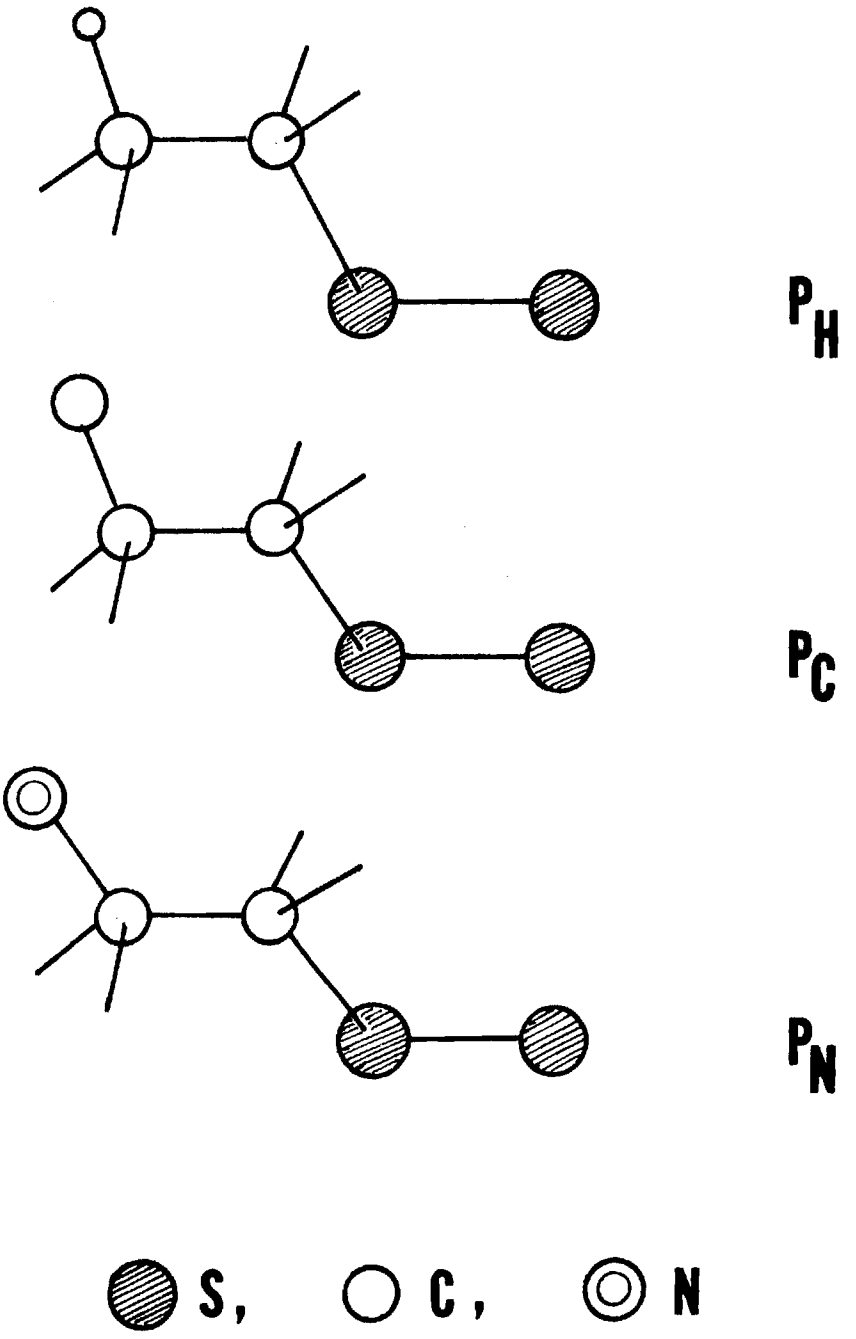

Figure 6. Various Conformations of Half-Cystine Group

X = carbon (P_C) and at 700 cm^{-1} for X = nitrogen (P_N). The schematic diagram for these types is shown in Figure 6. Miyazawa and Sugeta (1974) assigned 670 cm^{-1} line to one P_H form, 724 cm^{-1} to five P_C forms and 700 cm^{-1} to two P_N forms. The author examined a model of the three dimensional structure of Phillips (1966) and observed one P_H form of residue 94, five P_C forms of residues 6, 30, 80, 115, 127 and two P_N forms of residues 64, 76. However, the C-S stretching from methionine group was also expected around 700 cm^{-1} region (Nogami et al. 1975, Miyazawa and Sugeta, 1974). The correlation between the C-S stretching frequencies and internal rotation about the CH_2-CH_2 and CH_2-S bonds was made as follows: 760 and 719 cm^{-1} for the trans-trans form, 746 and 697 cm^{-1} for the trans-gauch form, 667 for the gauch-trans form and 723 and 645 cm^{-1} for the gauch-gauch form. Therefore, it was suggested that the line near 700 cm^{-1} in the spectrum of lysozyme crystal is partially contributed by the two methionine groups in the trans-gauch form. From the solution spectrum in Figure 2(b) and the powder spectrum in Figure 3, no noticeable changes are observed in the lines arising from P_H, P_N and P_C forms indicating no conformational variation. But the Figure 4 shows a complete destabilization effect of TCA on this region. The explanation for the decrease in intensity of those corresponding peaks can not be made at present stage. But it may be possible to assume that the C-S stretching vibration in protein molecules may

be affected not only by conformational factor but also local environmental change of the corresponding groups. Because, only decrease in the intensity of Raman lines around 700 cm^{-1} region is observed without any appearance of new or strong peaks.

One of the most sensitive regions to structural changes are those bands in region of the C-C and C-N skeletal stretching vibration. There are marked changes in shape noticed in the spectra of solution, powder and TCA-treated solid. Sensitivity of these bands were also observed by Brunner and Sussner (1972) in study of heat denaturation and by Chen et al. (1974) in chemical denaturation. But further studies are needed for explanation in detail.

In addition, the author presented a spectrum of TCA-treated solution of lysozyme in Figure 7 to check whether its destabilization process is reversible or irreversible. The spectrum shows the same feature as that of native lysozyme solution indicating that the denaturation is reversible and that the TCA precipitating method widely used in protein isolation may be acceptable.

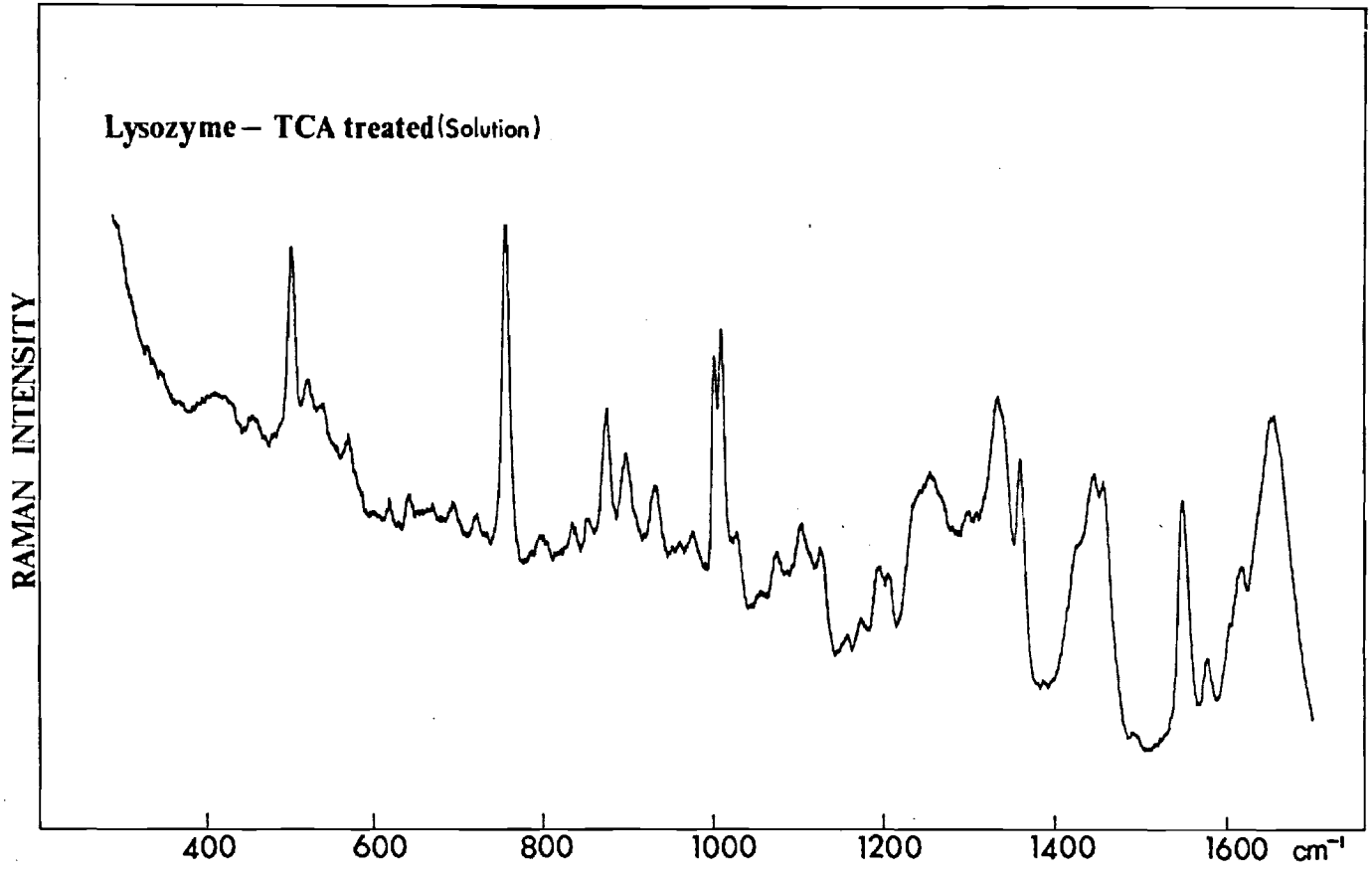

Figure 7. Raman Spectrum of TCA-treated Lysozyme in Solution

CHAPTER II

STRUCTURAL VARIATIONS OF RIBONUCLEASE A(RNase A) AND CARBOXYPEPTIDASE A(CPDase A)

Introduction

For continuing studies of Chapter I of this dissertation, the author wishes to extend Raman spectroscopic experiments further to the model systems, ribonuclease A (RNase A) and carboxypeptidase A(CPDase A).

Bovine pancreatic RNase A is a chain-cutting enzyme with a molecular weight of 13,700 which is specific for cleaving the P-O bond on the far side of a phosphorus which is connected to the 3' carbon of a ribose ring that bears a pyrimidine (C or U). The primary sequence of its 124 component amino acids is cross-linked by four disulfide bridges.

The author has carried out systematic Raman spectroscopic studies of RNase A under various conditions. This is an attempt to explore further the relationship between Raman bands and protein structures. First, the author wishes to show that the effect of solvation on the conformation of the protein can be detected by examining the features of laser Raman bands. The Raman spectra of RNase A in solution were first obtained and interpreted by Lord and Yu (1970b).

But the interpretation was mostly limited to no more than the identification of the amino acid constituents. The powder spectra were not reported at that time. Recently, Koenig (1972) obtained a preliminary Raman spectrum of the same protein in the solid state and compared it to that in solution. He noted that the Raman lines of the aqueous solution spectra are sharper than the lines of the solid spectra. In Figure 1, the author presents the higher resolution Raman spectra of RNase A in the powder form and in solution. Two powder spectra were reported with one taken in 0% and the other in 100% relative humidity at 25°C. For the direct comparison between the solid and solution phases, extra spectra are presented in Figure 2. The author then proceeded to study the pH dependence of the protein conformation. The question concerning the validity for the structural comparison of the protein between crystal and solution states should be raised in the RNase A case. Therefore, Raman spectra of the single-crystal and crystalline powder of RNase A were obtained for comparison with that of solution. Also the TCA effect on RNase A was dectected by Raman spectroscopy as it had been for the case of lysozyme described in Chapter I.

In addition, the Raman spectrum of carboxypeptidase A (CPDase A) (Anson) crystals was compared to that of solution and shows that there exists a small difference in the line shape of the amide III region. Recently, the assumption

associated with functional aspects with respect to the crystalline state was questioned for CPDase A by Johansen and Valee (1971). By treating CPDase crystals with diazotized p-arsanilic acid to modify tyrosine-248, an active-site residue of the enzyme, and comparing the visible spectra of crystals and solution, they concluded that the arsanilazocarboxypeptidase enzyme changed the side-chain conformation when it was crystallized.

CPDase A is an enzyme that digests polypeptide chains from the carboxyl terminal end. The molecule is composed of 307 residues joined together in a polypeptide chain having one disulfide bond at positions 138 and 161 and a molecular wieght of 34,489 (Pétra 1970). CPDase A is an unusual protein which contains appreciable amounts of parallel-β-pleated sheet structure. Raman spectra of such a protein have not been reported so far in the literature. A spectrum of the heat denatured protein was obtaned to see if the rigidity of the protein is maintained by the disulfide bond.

Materials and Methods

Lyophilized powder of beef pancreas RNase A (Lot No. RAF IFC) were purchased from the Worthington Biochemical Corp. and used without further purification. The sampel in powder form was packed into a conical depression at the end of an 1/8-in. gold plated copper rod, which was fastened in a Thermovac flask equipped with an "O" ring and a vacuum

tight stopcock. The powder was dried in vacuo at 25°C over phosphorus pentoxide for 2 hours before the experiment. The flask containing the sample and phosphorus pentoxide was then filled with dry nitrogen at 1 atm and used for laser Raman scattering work.

Lyophilized powder of RNase A in 100% relative humidity was prepared by equilibrating the powder at room temperature with saturated H_2O vapor pressure (about 24.00 mmH_g) for 5 hours.

The solution sample at 200 mg/ml concentration and pH 8.89 was obtained by dissolving the required amount of the powder used for solid sample in doubly distilled and degassed water without adding HCl or NaOH solution. pH adjustment ofr the samples of pH 5.00 and pH 1.70 was made with 0.1 N HCl solution.

L-Tyrosine hydrochloride in the solid form was purchased from the Sigma Chemical Co. (Cat. No. T-2006) and used without further purification and the concentration of the solution sample prepared from the powder was 0.1 M and the pH value 1.0.

Crystals of beef pancreas RNase A grown from 55% 2-methyl-2,4-pentanediol(MPD) water solution at pH 5.0 (Kartha et al. 1967) was a gift from Dr. J. Bello of Roswell Park Memorial Institute, Buffalo, N. Y.. They were of assorted shapes and sizes. The maximum dimension of some crystals was as large as approximately 3 mm. These crystals

were kept in 75% MPD when the author received them. Another type of RNase A crystals was prepared from ethanol solution according to the method of McDonald (1955). 100 mg of RNase powder were dissolved in 0.5 ml of H_2O and cooled to $5^{\circ}C$, and 6 ml of 95% ethanol of the same temperature was added with stirring. A heavy amorphous precipitate was formed and left on standing at 10 to $20^{\circ}C$ for several hours. This preparation produced a mass of fan-shaped rosettes of rectangular or needle-shaped crystals.

The preparation of TCA-treated RNase A powder followed the same way as described in Chapter I.

Crystals of beef pancrease CPDase A (Anson) were purchased from the Worthington Biochemical Corp. (Worthington Lot No. COA-1KB). The solution sample of CPDase A was prepared by dissolving a 30 mg sample in 0.5 ml of 3 M NaCl solution at pH 7.00, since CPDase A is not soluble at neutral pH and low ionic strength. The sample of heat denatured CPDase A was prepared by diluting the above solution sample ten times to 6 mg/ml at pH 7.00 and heating at $100^{\circ}C$ for 45 minutes. The solid sample was obtained by air-drying the resulting precipitate.

The sample of neurotoxin α from Egyptian Cobra (Naja haje haje) was a gift from Dr. D. J. Strydom of National Chemical Research Laboratory, Republic of South Africa. Cobramines A and B were kindly supplied by Dr. J. Wolff of National Institute of Arthritis, Metabolism, and Digestive

Disease.

For Raman scattering experiments, a large single crystal of RNase A was transferred to a 5 mm diameter glass vial filled with 75% MPD. The laser beam was then sent through the crystal in the vial and the scattered light directly from the scattering column inside the crystal was collected by an F/1.1 lens on the entrance slit of a Spex 1401 double monochromator. The shape of the crystal remained unchanged after long exposure (about 18 hours) to the laser beam. After completion of the experiments with the single crystal at various orientations, the crystal was crushed into finely divided particles. The sizes of the crystals were much smaller than the diameter of the focused laser beam (about 50 μ). The Raman spectra were then obtained with the crystalline powder precipitated at the bottom of a 1.0 mm diameter capillary containing 75% MPD. The thickness of the powder layer was about 0.5 mm. This capillary cell was held vertically by a Cary instrument's kinematic base and backing mirror assembly. The laser beam entered the cell from the bottom and traveled through the layer. The light scattered at 90° to the incident beam was then collected and analyzed. For the spectra of RNase A crystalline powder from ethanol solution, the same method as described in Chapter I for the crystalline lysozyme powder was followed except that the sample was equilibrated at room temperature with the vapor of 88% EtOH solution for 12 hours.

Raman spectra of CPDase (Anson) crystals were obtained by the same procedure as that for the crystalline lysozyme powder in Chapter I. Before the experiments, crystals of CPDase A were washed with water three times by centrifugation and decantation of the supernatant solution. The handling techniques for powder samples used in these experiments were similar to those described in Chapter I.

All the spectra reported here were obtained with the 514.5 nm line of an argon ion laser and at 4 cm^{-1} resolution. The Raman system and spectroscopic method was described in Chapter I. But the spectra in Figure 12 for comparison between native crystals of CPDase A and its heat denatured powder were recorded on a Spex Ramalog system with triple monochromator (1401 double monochromator plus third monochromator) equipped with a RCA model C-31034 phototube and a modified Spex pc-1 photon counting system.

Results and Discussion

Comparison of Ribonuclease A(RNase A) Structures in Various States

Effects of Humidity and Solvation on the Conformation of RNase A. In Figure 1, there are presented the higher resolution Raman spectra of RNase A in the powder form and in solution. Between the two powder spectra which are taken in 0% relative humidity for one and 100% for the other, no noticeable spectral differences were observed. On comparing the powder and solution spectra, however, several regions

have revealed striking changes.

The effect of solvation on the spectrum of RNase A powder is quite striking, particularly in the region of 500-870 cm^{-1} and the amide III backbone. The most dramatic change is the intensity decrease of the lines at 644 and 854 cm^{-1} relative to the line at 834 cm^{-1}. These three lines have been assigned to the ring vibrations of tyrosine (Lord and Yu 1970b). In the S-S and C-S stretching region (500-725 cm^{-1}), the half-width of the S-S line at 516 cm^{-1} decreased from about 20 cm^{-1} to 15 cm^{-1} and marked changes in intensity at 657 and 724 cm^{-1} are noticed. In the amide III region, the two lines at 1239 (β structure) and 1260 cm^{-1} (α helix) have sharpened and the latter shifted to 1265 cm^{-1}. Preliminary interpretations of these significant spectral changes were given earlier (Yu et al. 1972).

Upon solvation, the intensities of the lines at 644 and 852 cm^{-1} have decreased considerably relative to the line at 832 cm^{-1}. Since three of the six tyrosines (Tyr 73, 76 and 115) have a normal pK_a value (10-10.2) (Shugar 1952; Woody et al. 1966, Tanford and Hauenstein 1956, Tanford et al. 1955) and are presumably accessible to the water molecules (Wyckoff et al. 1967), one might interpret these intensity changes as due to the solvation of these exposed tyrosines when the protein molecule passed into the solution and not necessarily the result of conformational changes. However, the author has interpreted the Raman

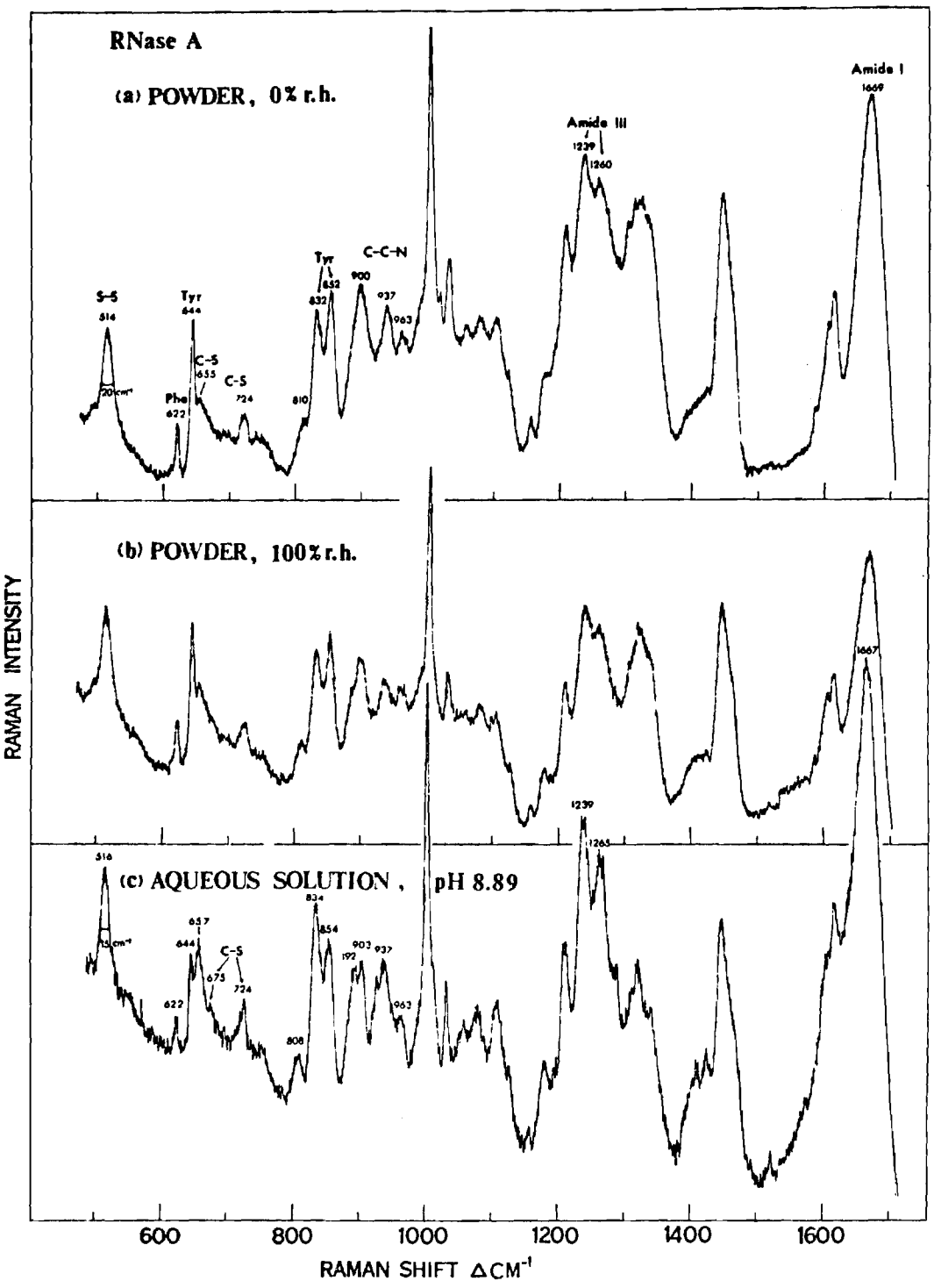

Figure 1. Raman Spectra of RNase A in Solid (0 and 100% r.h.) and Aqueous Solution (pH 8.89)

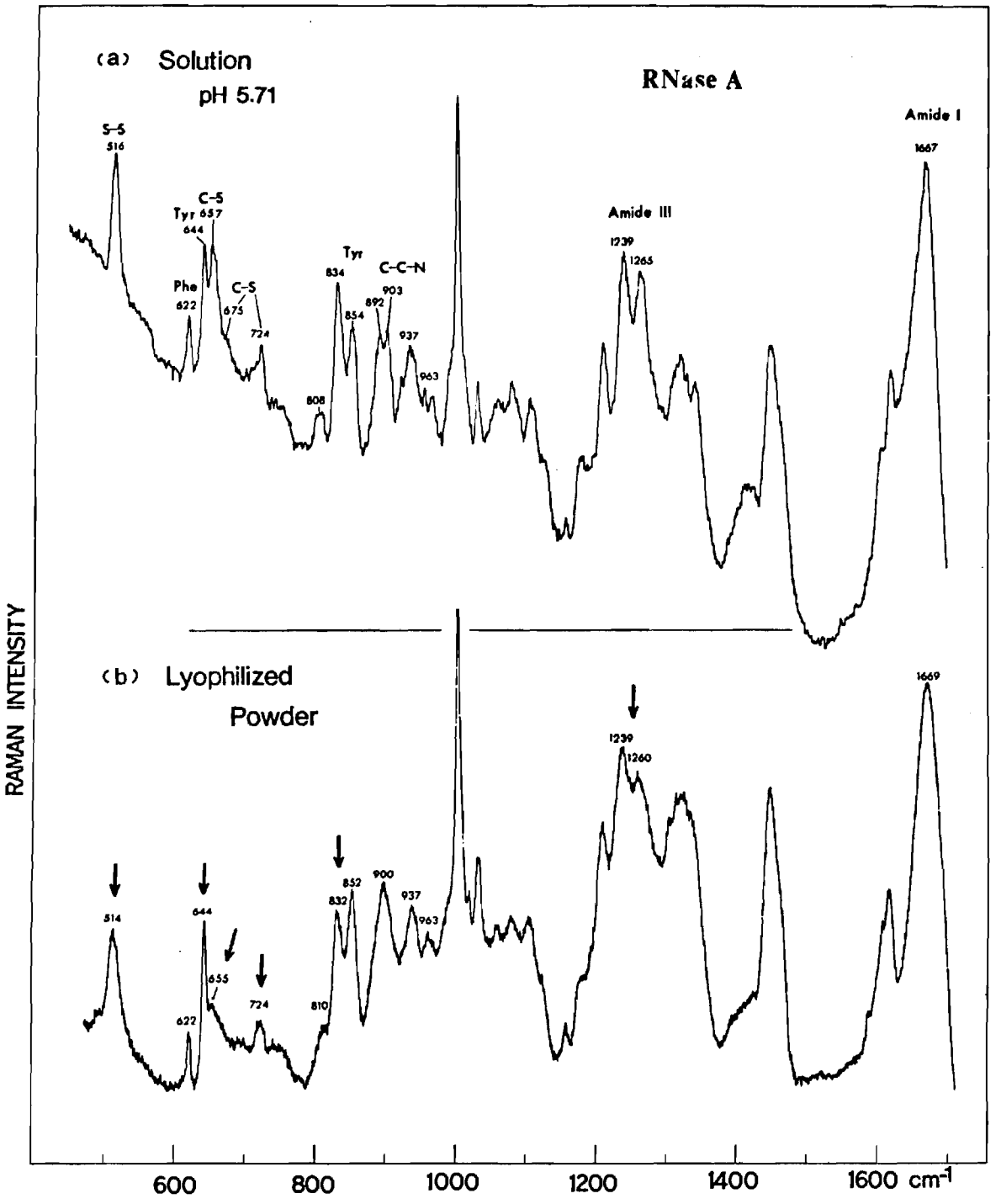

Figure 2. Raman Spectra of RNase A in Solution (pH 5.71) and Solid (0% r.h.)

intensity changes as due to the changes in the local environment of the three "buried" tyrosines (i.e., Tyr 25, 92 and 97) which probably is the result of conformational changes in the vicinity of these groups. This interpretation is primarily based on the fact that the direction of the relative intensity changes at 644, 834 and 854 cm^{-1} upon dissolution of RNase A was opposite to that observed for amino acid tyrosine on dissolution as shown in Figure 3.

The author thought that more convincing evidence might be obtained by examining the Raman spectrum of a protein solution whose tyrosine(s) is(are) known to be "buried". The author has found that neurotoxin α is a good model system. This toxin protein from Naja haje haje venom consists of 61 amino acid residues and four disulfide bridges (Botes and Strydom 1969). Its single tyrosine-24 is known to have an abnormally high pK_a value (i.e. 11.9) (Chicheportiche et al. 1972) and is believed to be "buried". The spectra of solution and solid of this toxin are shown in Figure 4. As expected the lines at 644 and 846 cm^{-1} are smaller for solution than for solid. This has led the author to believe that the intensity decreases of the tyrosine lines at 644 and 854 cm^{-1} in RNase A spectra are the manifestation of conformational changes in the vicinity of those "buried" tyrosines.

Additional evidence in support of the above interpretation comes from the Raman spectra of cobramines A and

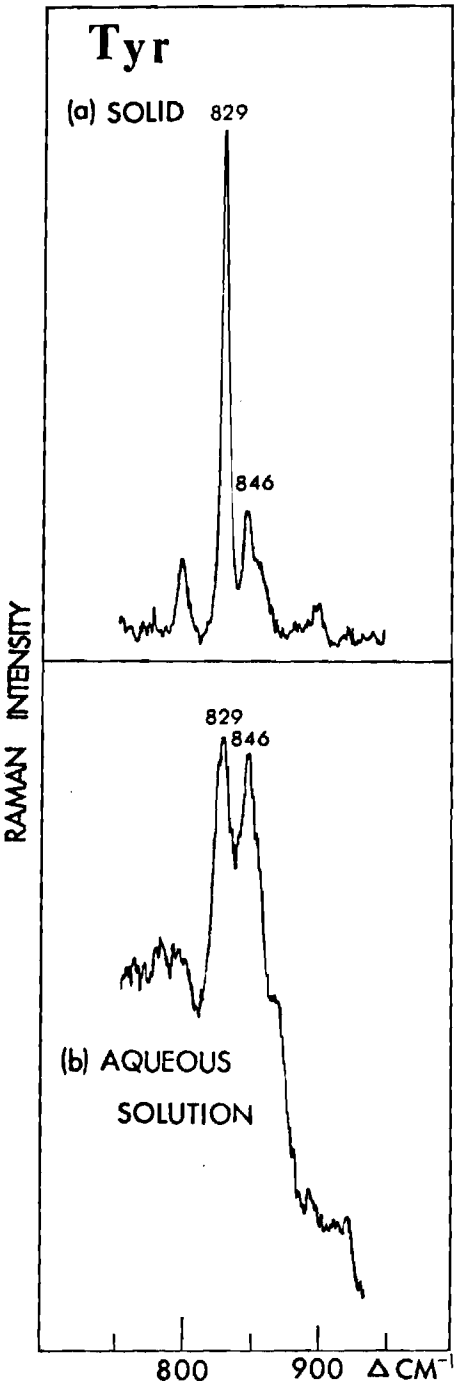

Figure 3. Raman Spectra of Tyrosine in Solid and Aqueous Solution

B. These two are heat-stable, basic proteins from India cobra venom, containing four and three tyrosines, respectively (Hayashie et al. 1971, Larsen and Wolff 1968). Recent studies (Yu et al. 1973) on the temperature variation of the intensity ratio of the 644, 830 and 853 cm^{-1} lines in cobramine B suggested that all three tyrosines were "buried" and probably involved in interactions similar to those of the three "buried" tyrosines in RNase A. Although, at present, the status of four tyrosines in cobramine A is not known, based on the intensity ratio argument used by Yu et al. (1973) in their studies of cobramine B, it is quite likely that all tyrosines in cobramine A are also "buried". The Raman spectra of these two proteins in the 600-870 cm^{-1} regions are shown in Figure 4.

Small but definite spectral changes occur in the S-S and C-S stretching vibration region (500-730 cm^{-1}). The lines at 514 and 655 cm^{-1} were assigned to the S-S and C-S stretching of the four disulfide bridges and the intensity ratio of the C-S to S-S lines was related to the C-S-S bond angles by Lord and Yu (1970a, b). But recent studies of model compounds (Sugeta et al. 1972, 1973, Miyazawa and Sugeta 1974) revealed that the C-S stretching vibration depends on the rotational conformation of the C-C bond adjacent to the C-S bond. This explanation may be interpreted to mean that there is no relation between the S-S and C-S stretching lines. They found that the Raman line at

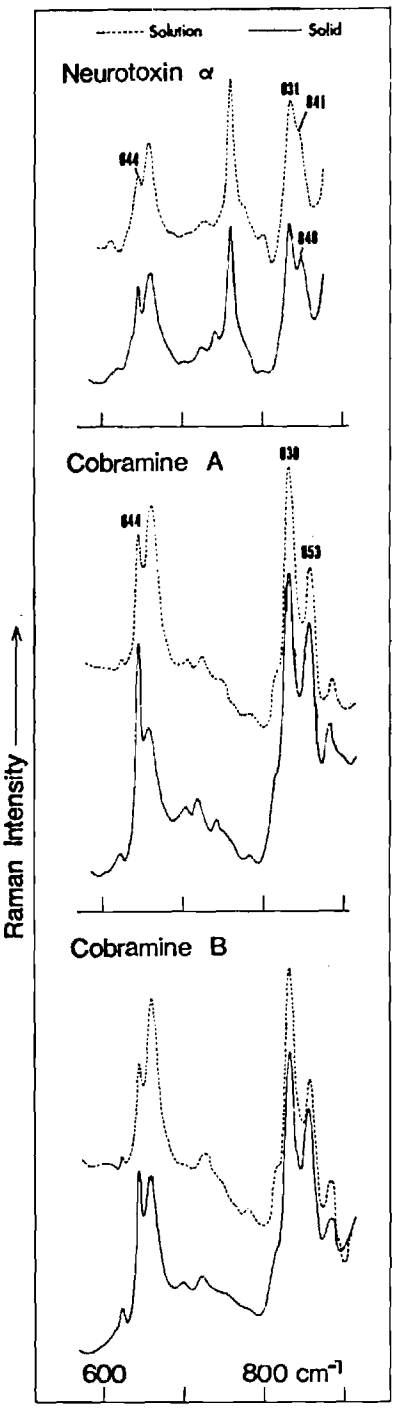

Figure 4. Raman Spectra of Tyrosyl Groups in Various Model Compounds

510 cm^{-1} appeared for dialkyl disulfide compounds whose conformation about S-S bonds is in the gauch-gauch-gauch form and assigned the 516 cm^{-1} line observed in RNase A solution spectrum to the gauch-gauch-gauch form. According to the X-ray studies of Wyckoff et al. (1970), four cystine groups in the molecule of RNase in crystal are in the gauch-gauch-gauch form and all the eight half-cystine groups are in the P_C form (for notation, refer to Chapter I of this dissertation). The C-S stretching vibration of the P_C form was observed in the 700-750 cm^{-1} region by Sugeta et al. (1972, 1973) in their studies of model compounds. On this basis, they assigned the line 725 cm^{-1} to the C-S stretching vibration of the P_C form of half-cystine groups and one trans-trans form of methionine group (residue 79). Their analysis of C-S stretching frequencies of the methionine group was mentioned in Chapter I. The line at 657 cm^{-1} which was assigned to the C-S stretching vibration of the cystine groups by Lord and Yu (1970a, b) earlier has been reassigned to the C-S stretching vibration of two methionine groups in the gauch-gauch form (residues 29, 30). The peak for one methionine in the trans-gauch was not clearly shown up but still observable at 700 cm^{-1} as a shoulder.

In the powder spectrum, the half-width of the 514 cm^{-1} line is 20 cm^{-1}, while it is only 15 cm^{-1} in the solution spectrum. This may reflect that the S-S bond is more

uniform in solution than in the solid state. The explanation for the striking increase in intensity of the line at 657 cm^{-1} on dissolution may be speculated in two ways. First, it might be assumed that a number of the P_C form transformed into the P_H form. However, this is not an acceptable explanation, because there is no decrease in intensity of the 724 cm^{-1} which is also responsible for the P_C form. The other speculation is that either more gauch-gauch form or gauch-trans form (~667 cm^{-1} expected) of the methionine groups are produced on dissolution. Then, there should be some change in intensity of the lines at 697 and 746 cm^{-1} (trans-gauch) or 719 and 760 cm^{-1} (trans-trans). Although there seems a subtle decrease occurrs in 701 cm^{-1} region, but it does not look sufficient to support the second speculation. The peak of 657 cm^{-1} line in solution seems too strong to be explained so simply by the second speculation. Although further study on this subject should be made to elucidate the intensity variation of this characteristic line, the author tentatively interprets it as due to changes in the local environment of C-S bond of methionine groups in the gauch-gauch form.

The backbone conformation of RNase A is best assessed in the amide III vibration region (1220-1300 cm^{-1}). Two resolved lines are observed at 1239 and 1260 cm^{-1} in the powder spectra. According to a recent study on the Raman spectra of glucagon in various conformational states, Yu and

Liu (1972) concluded that the α helical, random coil and β structure of a protein should have the amide III vibrational frequencies at 1266, 1248 and 1232 cm^{-1} respectively. On the other hand, the X-ray diffraction studies of RNase A (Kartha et al. 1967) and RNase S (Wyckoff et al. 1970) have revealed that the main chain of the molecule may roughly be divided into two parts with three sections of α helix (2 to 12, 26 to 33 and 50 to 58, totaling approximately 15% of molecule) in the first half, and the bulk of the antiparallel-β structure (somewhat irregular) in the second half. With this information at hand, the author assigned the line at 1239 cm^{-1} to the antiparallel-β structure and one at 1260 cm^{-1} to the α helices. Upon dissolution, the 1260 cm^{-1} line shifted to 1265 cm^{-1} and both lines sharpened. Changes in frequency and line width in this region are considered a reflection of the backbone conformational changes indicating that the β structure and α helix parts of the molecule are more uniform in hydrogen bonds in the solution state.

In principle, the amide I region (1630-1700 cm^{-1}) should carry the same conformational information of a protein backbone. However, due to the fact that this region is less sensitive, the antiparallel-β structure and α helix components cannot be resolved and thus show up as a broad line at 1669 cm^{-1} in powder and at 1667 cm^{-1} in solution. The difference in the amide I frequency between

the spectra of solid and solution is within experimental uncertainty and considered insignificant.

The spectra of pH dependence of RNase A on conformation are presented in Figures 1, 2 and 5. There are no significant differences between the spectra of pH 8.89 and 5.00. But between the pH 5.00 and 1.70 spectra, there are several regions showing the spectral differences which are indicated by Arabic numbers in Figure 5. The residues which may respond to the pH change from 5.00 to 1.70 may include aspartic and glutamic acids whose side-chain contains carboxyl ion and the end carboxyl group of the protein. The aspartic acids whose carboxyl ion involves in hydrogen bond in crystal are Asp 14, Asp 83 and Asp 121 (Wyckoff et al. 1970). The carboxyl ion of Asp 14 forms a hydrogen bond to Tyr 25, Asp 83 to the NH of Cys 84 backbone and Asp 121 to the NH of Lys 66 backbone. Only one glutamic acid (Glu 111) participates in forming a hydrogen bond, which is with the NH_2 of Asn 111. On pH reduction to 1.70, protonation of the carboxyl groups of aspartic and glutamic acids may disrupt the hydrogen bonds formed by these carboxyl ions. This may cause the partial unfolding of the main chain as evident from spectral change in the amide III region, relative intensity changes of the Raman lines at 644 and 852 cm^{-1} to the 832 cm^{-1} line and changes in the C-S stretching vibration region.

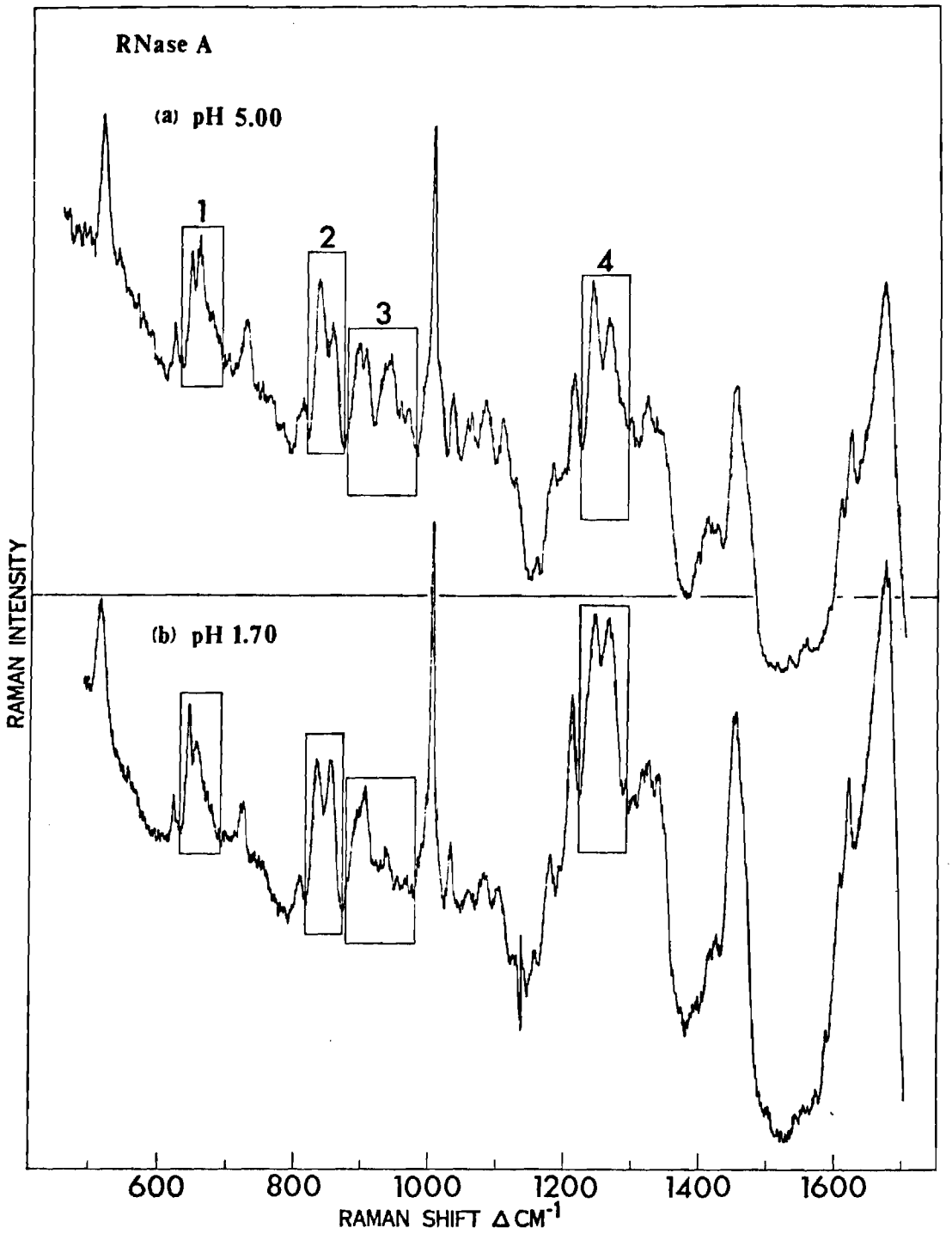

Figure 5. Raman Spectra of Aqueous RNase A at Different pH Values

Effect of Crystallization on the Conformation of RNase A. Here, the author reports the first Raman spectrum of a protein single crystal, RNase A, in Figure 6. Since the crystal contains about 40% of solvent and 60% protein by weight, the scattering intensities of many Raman lines from MPD are quite strong. However, there are two regions which are relatively clear and appear to be useful for the present studies: i.e., the 600-750 cm^{-1} and the amide III regions. In Figure 7, the author compares the spectra of solution, single crystal, and lyophilized powder in these two regions. A comparison between solution and crystal in Figure 7, shows both similarities and differences. In the amide III backbone region, both frequencies and line shapes show a good agreement between two phases. The author therefore concludes that there is no detectable effect of crystallization on the main-chain conformation of RNase A. On the other hand, in the 600-750 cm^{-1} region, the tyrosyl line at 644 cm^{-1} increased its intensity upon crystallization. The spectral feature in this region of the spectra does not depend significantly on the crystal orientation relative to the laser beam (see Figure 8 a-f) and is similar to that of crystalline powder. Based on the evidence from Raman spectra of neurotoxin α and cobramine A and B as discussed in the previous section, the author believes that the intensity increase of the 644 cm^{-1} line upon crystallization is a reflection of side-chain conformational changes.

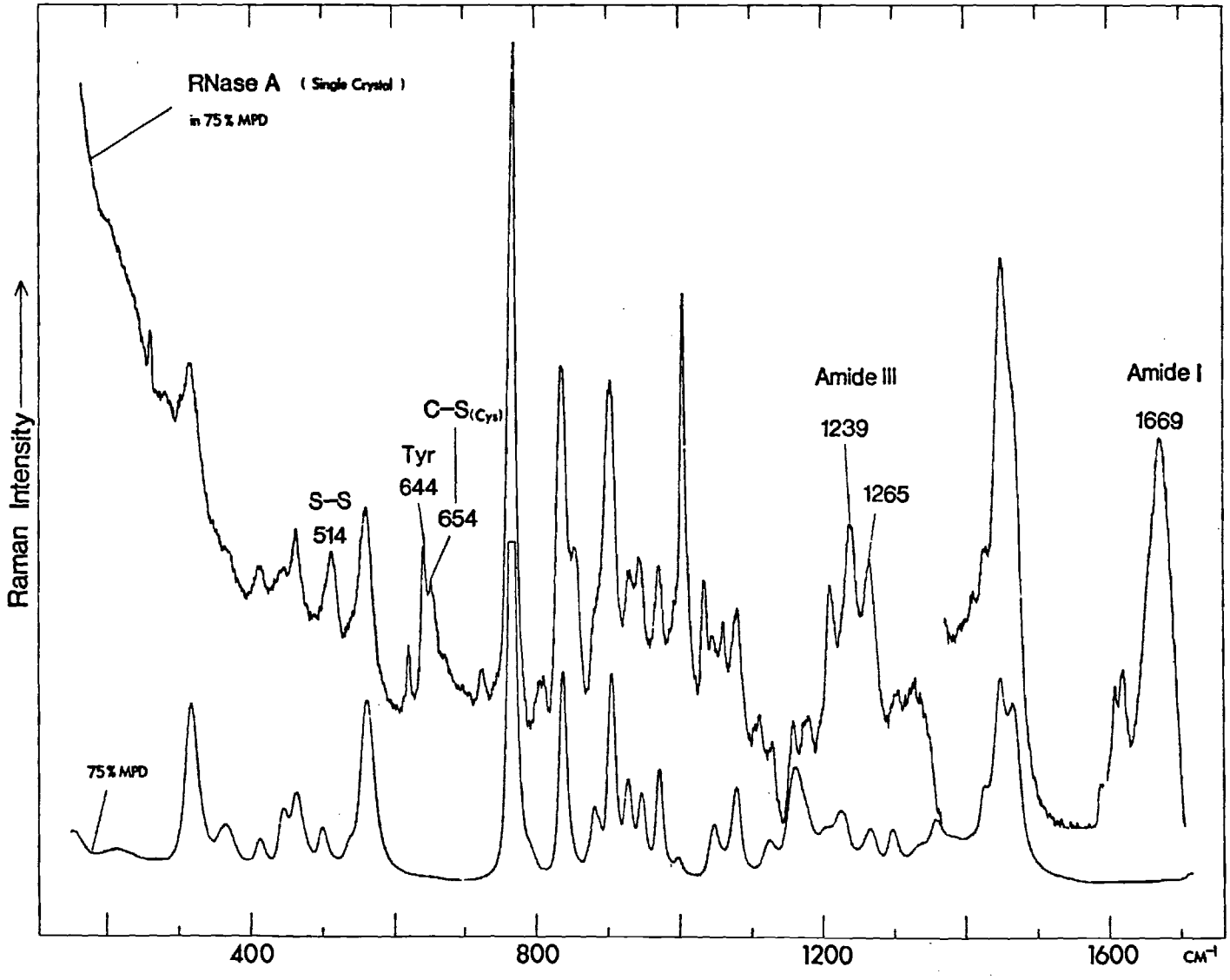

Figure 6. Raman Spectra of Single-Crystal RNase A and Mother Liquid (75% MPD)

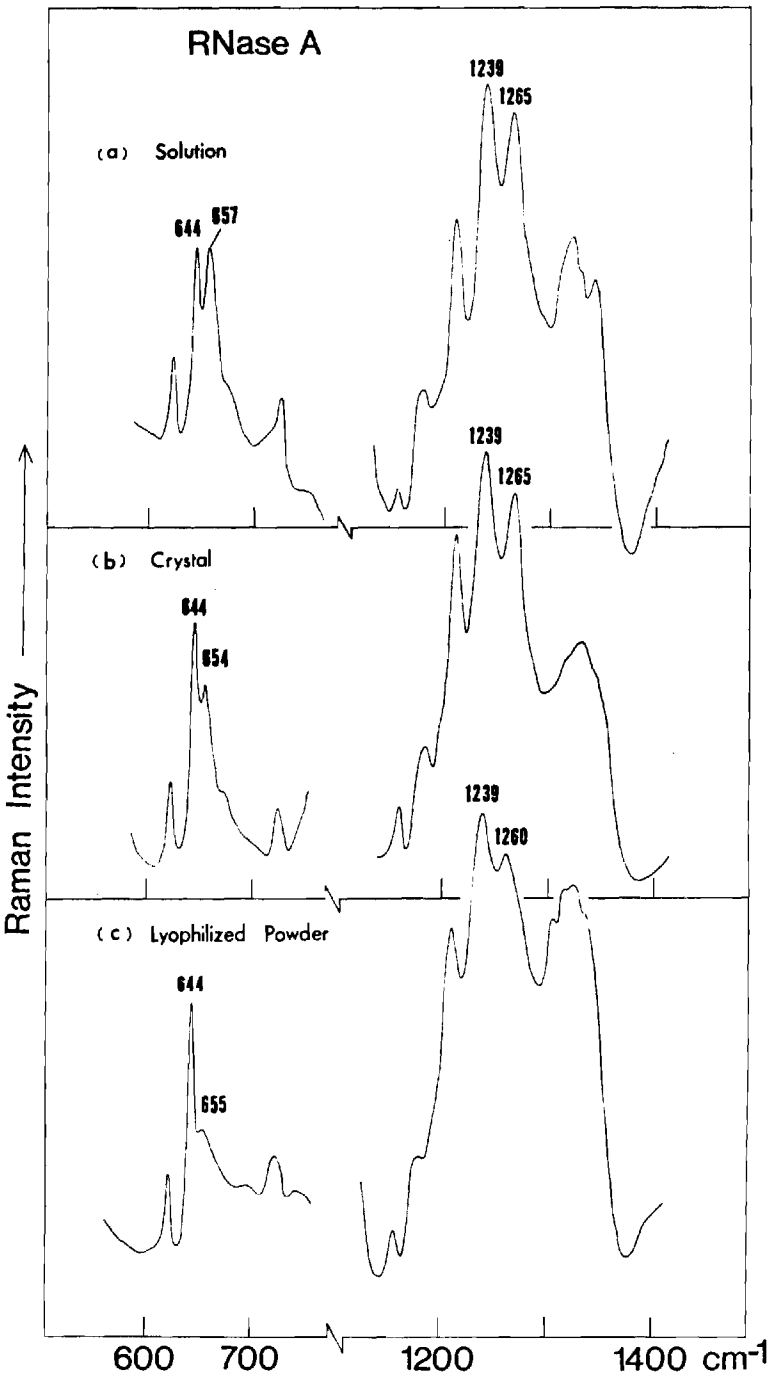

Figure 7. Comparison of Raman Spectra of Solution, Single-Crystal and Lyophilized RNase A

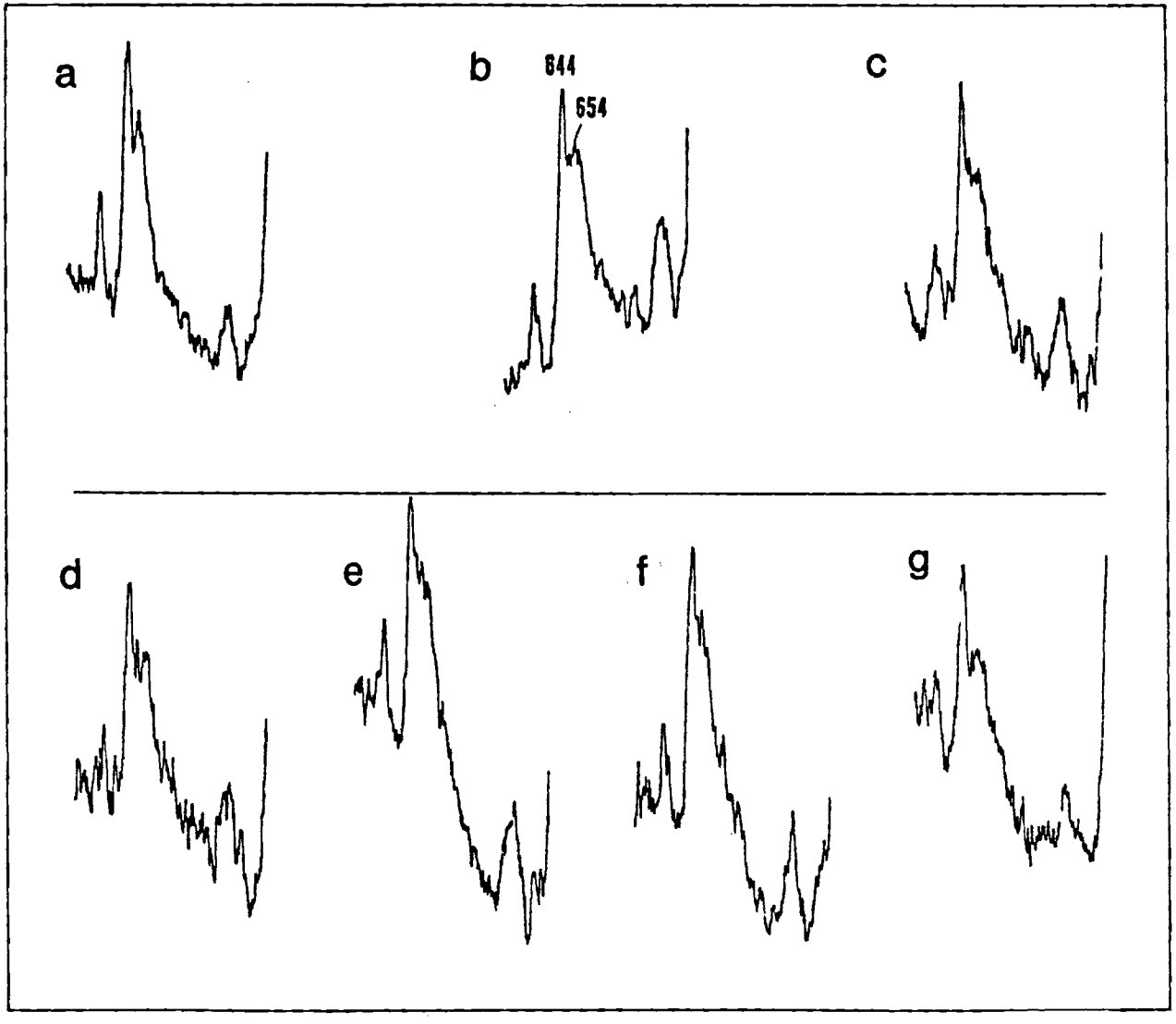

Figure 8. Raman Spectra of RNase A in Single-Crystal at Six Different Orientations

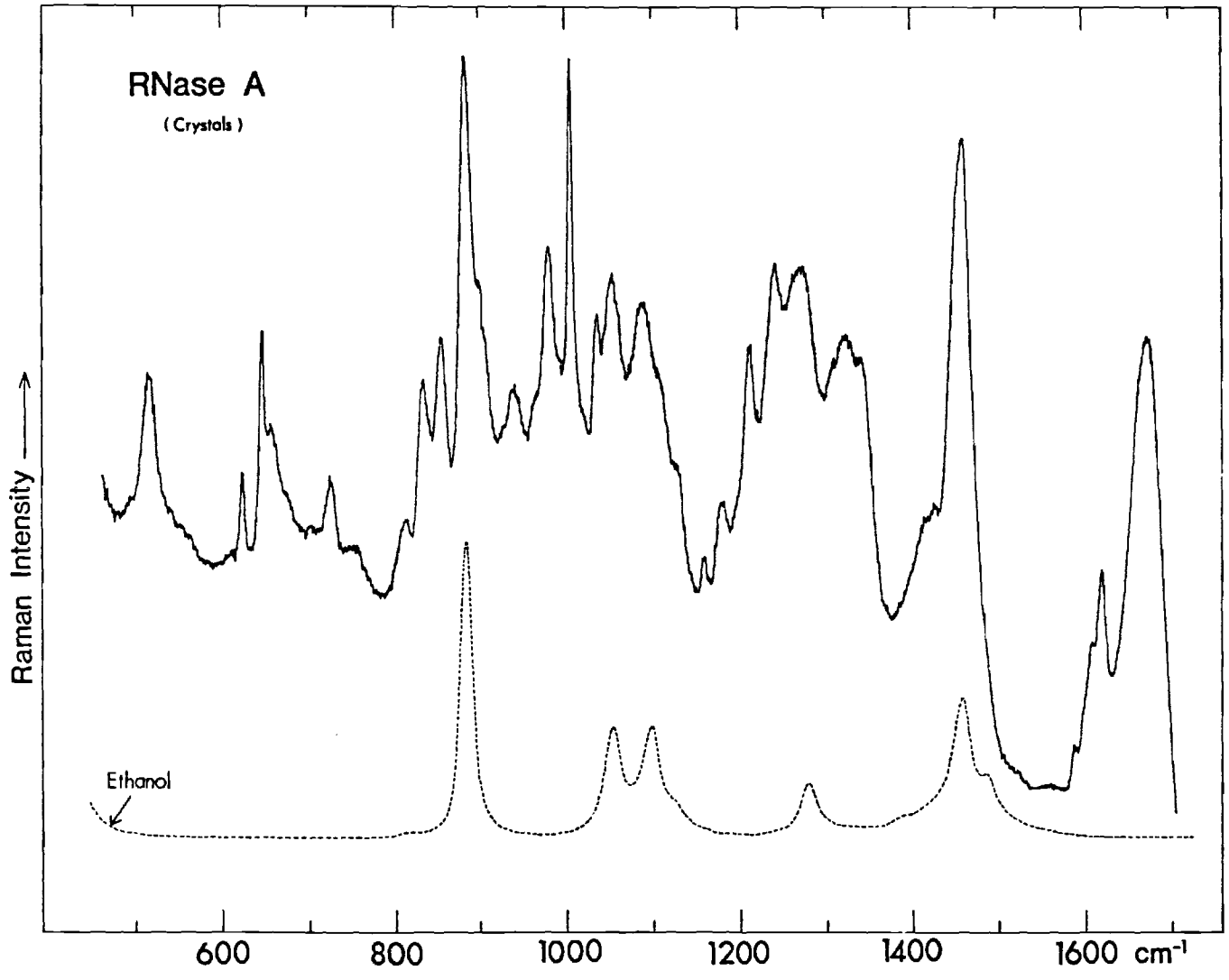

Figure 9. Raman Spectra of RNase A Crystalline Powder and Mother Liquid (88% EtOH)

An additional Raman spectrum of RNase A crystals is shown in Figure 9 (McDonald 1955). The RNase A crystalline powder was obtained from ethanol solution. The spectral feature in 600-750 cm^{-1} region is similar to that of RNase A crystals from MPD solution, indicating similar side-chain conformational changes.

Trichloroacetic Acid(TCA) Effect on the Conformation of RNase A. In Chapter I it was evidenced by Raman spectroscopy that TCA has marked effects of destabilization on the lysozyme conformation. Similar effects are observed in the spectrum of the TCA-treated RNase A as shown in Figure 10, particularly in the amide III and amide I regions. The sharpening of the amide III peak at 1236 cm^{-1} and the amide III peak at 1671 cm^{-1} is the same as that for the TCA-treated lysozyme reflecting that the main-chain backbone conformation of the protein becomes mostly antiparallel-β structure having more uniformed hydrogen bonds. The Raman line at 1260 cm^{-1} which reflects α helices has diminished. In the S-S stretching region of the disulfide frequency, there is a definite broadening and decrease in peak intensity and a frequency shift to 510 cm^{-1}. However, its area seems to be unchanged from that of the native enzyme. This means that the four disulfide bonds stay intact but their geometry is somewhat less uniform than that of the native enzyme. There is a weakening in peak intensity of the 655 cm^{-1} line which is accompanied by increase in intensity of

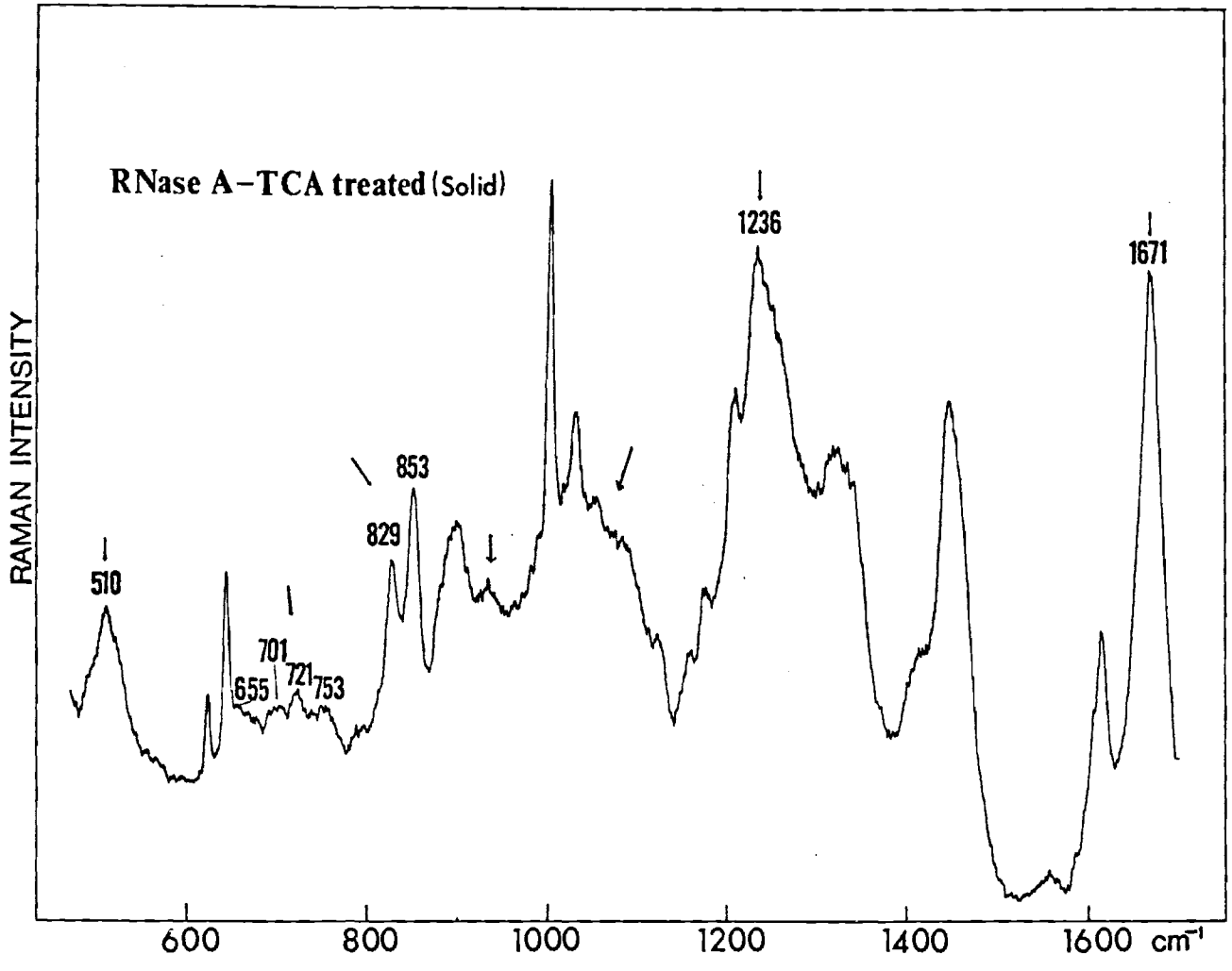

Figure 10. Raman Spectrum of TCA-treated RNase A Powder

the 701 and 753 cm^{-1} lines. It may reflect some changes in either conformation or local environment of the methionine side-chains. The ratio of relative intensities of tyrosyl groups, 852/832 cm^{-1}, is greater than that of the native powder. It may be an indication that the intra-interaction of "buried" tyrosyl groups becomes weaker. The definite broadening and weakening of the Raman bands in the C-C and C-N stretching vibration region is another indication of the backbone conformation change of the protein. The author checked whether the denaturation process is reversible or irreversible by examining the solution spectrum of TCA-treated RNase A and found that the spectrum features are the same as those of RNase A native solution. Same results were observed for the TCA-treated lysozyme as shown in Chapter I.

Conformation of Carboxypeptidase A(CPDase A)

Differences between the Backbone Conformation of CPDase A in the Crystalline State and in Solution. In the Raman studies of RNase A crystals, the author has just shown that there is no detectable difference between the backbone conformation of RNase A in the crystalline state and in solution. Also in Chapter I, the author has found that in the case of lysozyme the backbone conformation is the same as in solution. These conclusions have been based on the comparison of frequencies and line shapes in the amide III region of the Raman spectra. On the contrary, Yu et al. (1972), in their earlier studies of insulin crystals, have

observed significant spectral differences in the amide III region between the spectra of crystals and solution. These differences were interpreted in terms of backbone conformational differences between the two phases.

According to the X-ray analysis of CPDase A (Lipscomb et al. 1968, Quiocho and Lipscomb 1971), the structural heart of the molecule is a β sheet made up of no less than eight parallel or antiparallel extended chains (i.e., 4 pairs of parallel and 3 pairs of antiparallel chains) as shown in Figure 11. This massive pleated sheet covers roughly 20% of the total residues. Eight α helices which employ 35% of the total residues are packed on the two sides of the β sheet. The remaining residues occur in the "random-chain" (20%) and chain segments (25%) needed to tie these features together (Dickerson and Geis 1969).

The β sheet in CPDase A is constructed differently from that in lysozyme. The hydrogen bonds of the β sheet in CPDase A depend only upon amide-carbonyl bonds for its integrity, while the side-chains of serine, threonine, asparagine and glutamine are mostly responsible for the lysozyme β structure. In fact, most of the side-chains are hydrophobic in the β structure in CPDase A as would be expected in a β sheet designed for the interior of a molecule. The α helices have hydrophobic sides which are packed against the sheet. There are enough regular cross-chain hydrogen bonds that one can describe the structure as a true

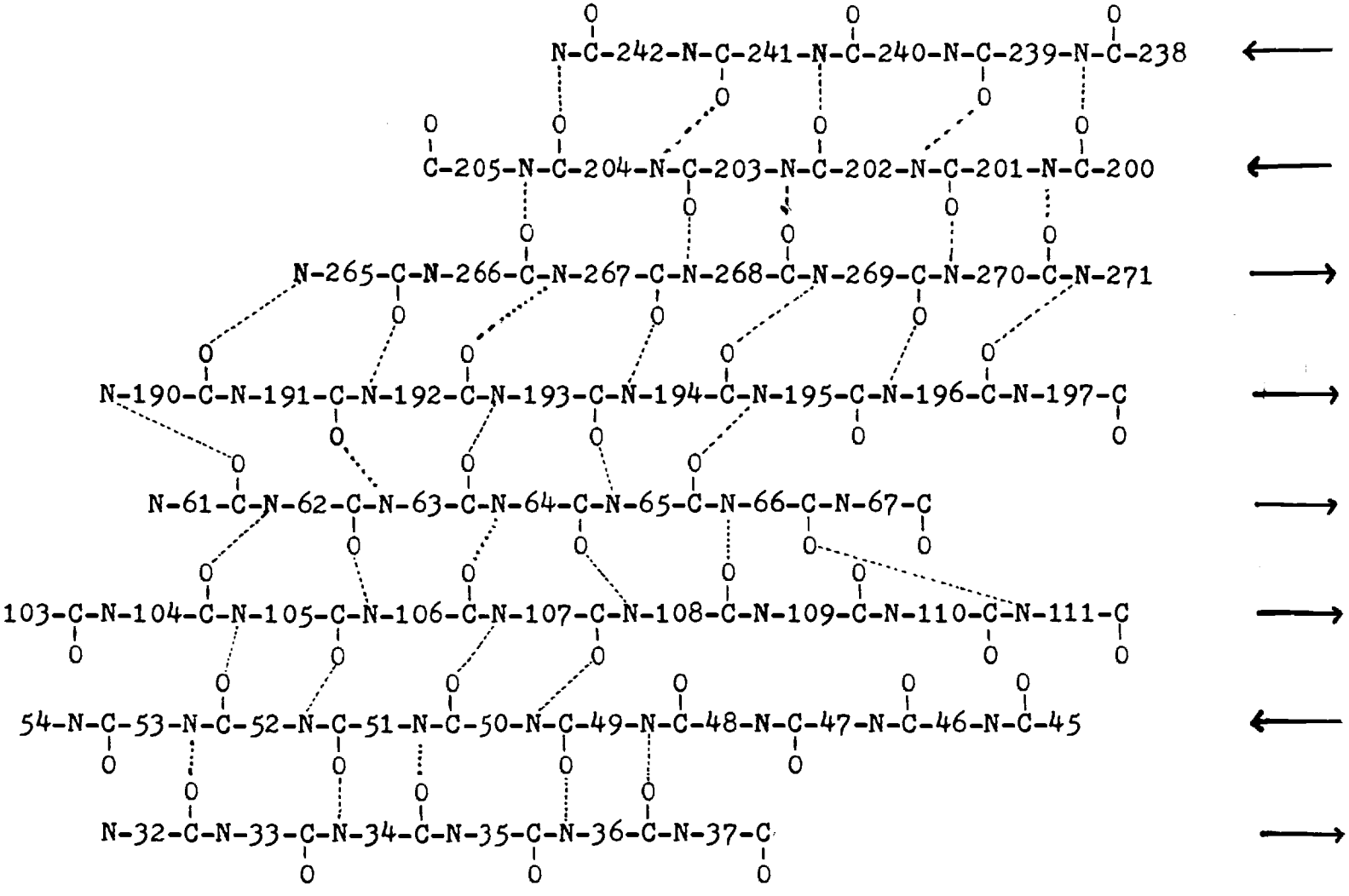

Figure 11. β Sheet of CPDase A (after Quiocho and Lipscomb 1971)

β-pleated sheet rather than merely "β-like folding".

Presently, the Raman spectral characterization of helix, random coil and antiparallel-β structure of protein is well established, but the data for parallel-β structure are still lacking. Raman spectra of CPDase A in crystals and in solution are compared in Figure 12. In crystals, two frequencies at 1247 and 1270 cm^{-1} may be identified as the amide III vibration. Upon dissolution, the broader component at 1247 cm^{-1} has become sharpened and the peak maximum was shifted to 1243 cm^{-1}. The author believes that the spectral changes observed here are the reflection of backbone conformational changes.

One unusual aspect of the spectral changes is the sharpening of the line at 1247 cm^{-1} on dissolution. Normally the opposite effect is observed with simple peptides (Koenig 1972). Although the exact nature of this line sharpening is not known, it might mean that the strength of intramolecular hydrogen bonds in solution is more uniform than in the crystalline state.

The assignment of the amide III lines in the spectra of CPDase A to various structural components of the backbone may be made on the basis of the criteria established by Yu and Liu (1972). In the Raman studies of conformational variations of glucagon, they concluded that the α-helical, random-coiled (H-bonded) and antiparallel-β structure of a protein should have the amide III vibrations near 1266, 1248

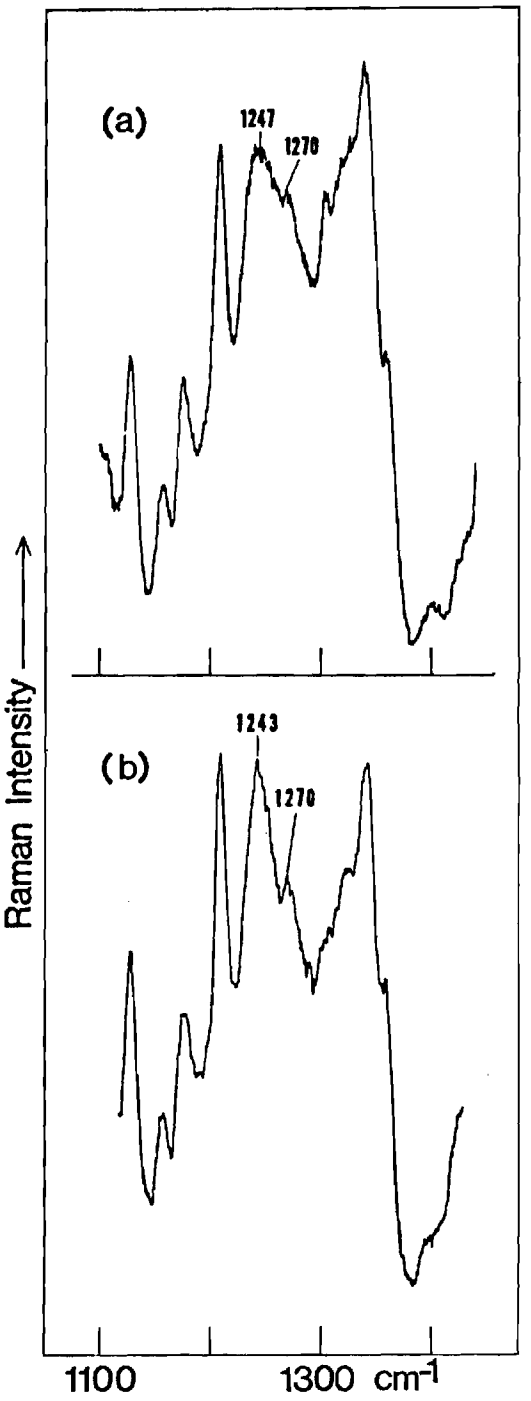

Figure 12. Raman Spectra of CPDase A (a) in Crystalline State and (b) in Solution

and 1232 cm^{-1}. Based on this information and structural knowledge from the X-ray data of CPDase A, it may be probable to assign the line at 1270 cm^{-1} to the α-helices and the one at 1247 cm^{-1} in crystals and 1243 cm^{-1} in solution to overlapping of parallel- and antiparallel-β-pleated sheet and random-coil structure. The line sharpening of the amide III band accompanied by a frequency shift to the lower value may indicate that the orderliness in hydrogen bonds of the backbone arises from the β-structure sheet. In order to confirm this assignment, Raman studies of model compounds known to exist predominantly in a parallel β structure seem desirable.

Comparison of Raman Spectra of Native and Heat Denatured CPDase A in the Solid State. The spectra of CPDase A in native crystalline powder and its heat denatured solid are shown in Figure 13 for comparison. The line at 523 cm^{-1} in crystals disappeared upon heat denaturation and no new line has come out near 500 cm^{-1}. The characteristic of this line is due to the S-S stretching vibration of the disulfide bond in the gauch-gauch-trans form. This assignment is based on the results of studies using model compounds by Sugeta et al. (1972, 1973). This conformational prediction about the disulfide bond is in good agreement with the structural model derived from X-ray data (Quiocho and Lipscomb 1971). CPDase A has only one disulfide bond connecting residues 138 and 161 which is located outside the molecule. Disappearance of this line upon denaturation may

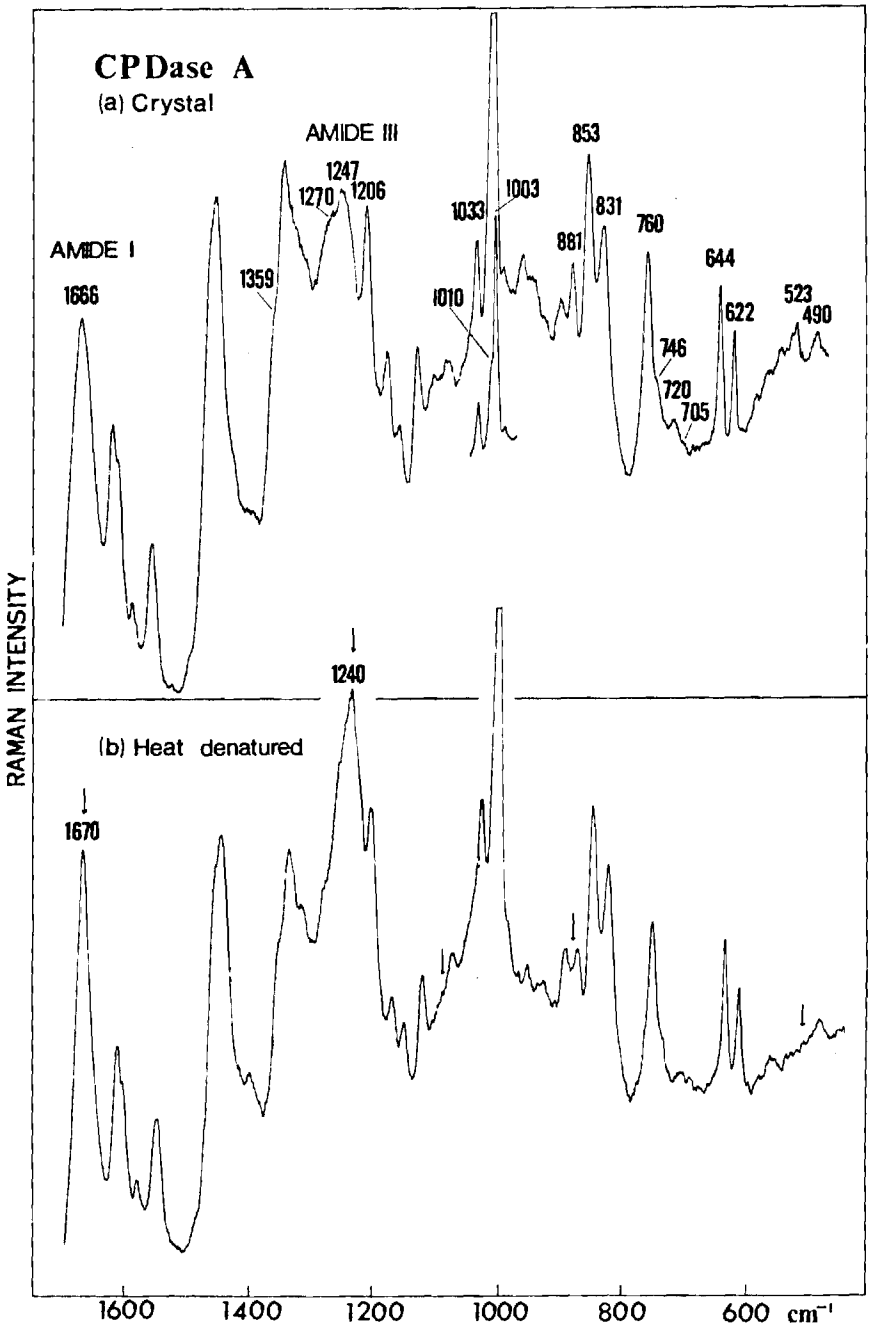

Figure 13. Raman Spectra of CPDase A (a) in Native Crystals and (b) in Heat-denatured Powder

be assumed that the disulfide bridge is broken.

Examination of the C-S stretching vibration region (650-750 cm^{-1}) gives information about the conformation of methionine and half-cystine side-chains. According to the three-dimensional model derived by X-ray studies (Quiocho and Lipscomb 1971), the molecule has two P_C forms of half-cystine groups and three methionine residues in which one (Met 22) is in the trans-trans form and two (Met 103 and Met 301) in the trans-gauch form. The author assigns the 720 cm^{-1} line in the native crystal spectrum to the C-S stretching vibrations of two half-cystine groups in the P_C form based on the studies of Sugeta et al. (1972, 1973). However, one methionine group in trans-trans may be partially responsible for that line (Nogami et al. 1975). The lines at 705 and 746 cm^{-1} are assigned to two methionine groups in the trans-gauch form. In the spectrum of heat-denatured CPDase A, subtle changes in shape and frequency of Raman bands are observed in this region. But the effect is too small to assume conformational changes about half-cystine and methionine side-chains. The observed variation of Raman lines may be attributed to no more than a subtle distortion of the bonds involved.

Quiocho and Lipscomb (1971) stated that their X-ray results concerning the location of tyrosine ring does not agree with chemical studies (Simpson and Valee 1966). According to the X-ray results, at least 16 out of 19

tyrosine OH groups are accessible to solvent while chemical conclusion claims that 7 to 8 tyrosines are at the surface and 11 or 12 tyrosines are located in the interior of the folded protein. From the present Raman spectrum of native CPDase A crystals and the Raman data of model compounds discussed for tyrosine side-chains in the previous section, X-ray results seem more reliable. But it must be mentioned that both the Raman data of CPDase A crystals and X-ray results are made from the solid state of the sample. As already evidenced by laser Raman studies, RNase A has different conformational situations of tyrosine groups in the solution and solid states. Therefore, a careful caution should be made when the protein structure is compared in solution and solid states. For the conformation of tyrosine groups in CPDase A molecule, it may not be proper at present to assess the result of X-ray with respect to the chemical studies. Raman data on CPDase A solution should be expected for making clear the argument. Upon heat-denaturation, the relative intensities of tyrosine lines at 644, 831 and 853 cm^{-1} do not change much indicating that thermal treatment does not affect the intramolecular interaction of the "buried" tyrosine rings. Of seven tryptophan residues of CPDase A (residue 151 is not included here since chemical sequence shows it as Phe (Bradshaw et al. 1969) while X-ray data define it as Trp (Quiocho and Lipscomb 1971)), five (73, 81, 126, 257, 294) are edge-exposed to solvent and two

(63, 147), quite buried (Quiocho and Lipscomb 1971). The relationship between Raman lines and tryptophan conformation were well discussed by Yu (1974) and Chen et al. (1973). Based on their discussion, the weak shoulder at 1359 cm^{-1} may be regarded as a reflection of most exposed tryptophan residues. As shown in the spectrum of heat-denatured CPDase A, the 1359 and 760 cm^{-1} lines arising from tryptophan ring vibration remain the same as that of native crystal spectrum. From the present data, it is hard to mention about the possibility of structural variation in two "buried" tryptophans. However, another region of 540-580 cm^{-1} which is also sensitive to tryptophan residues probably arising from in-plane bending vibrations shows some variation in intensity and band shape. It may be speculated as due to the effect of intermolecular interaction exerting on the tryptophan residues on the molecular surface which may be brought about by protein aggregation in the denaturing process.

There are serious variations noticed in the regions of C-C and C-N skeletal stretching vibration. It is thought that there must be a qualitative relation between the amide III and I bands of backbone conformation and these C-C-N stretching vibration region. This may be one of the subjects to be pursued in the future.

CHAPTER III

CONFORMATIONAL ANALYSIS OF SNAKE VENOM PROTEINS

Introduction

Of the many different kinds of animal venoms, snake venoms have been studied to the greatest extent, although not many studies are made concerning the physicochemical properties of its toxic components. The snake venom is classified into two groups, one is neurotoxin and the other, non-neurotoxin basic protein venom. The neurotoxin of the Elapidae family has been conveniently grouped depending on the number of constituent amino acid residues, Type I has 61-62 amino acid residues and Type II, 71-74. All the sea snake (Hydrophiidae) neurotoxins so far investigated belong to Type I. It is known that all of the Type I neurotoxins obtained from Elapidae and Hydrophiidae are similar regardless of species or geographic origin.

Cobramine B which belongs to the non-neurotoxic group is one of the two basic (isoelectric point 12.6) and heat stable proteins from the venoms of the Indian Cobra (Naja naja), which inhibits iodide accumulation by thyroid slices and have high cytotoxicity to Yoshida turmor cells (Larsen and Wolff 1968, Hayashi et al. 1971, Braganca et al. 1967).

Cobramine B consists of 52 amino acids having molecular weight of 5,840, of which six half-cystine residues form disulfide bridges (Larsen and Wolff 1968).

Neurotoxin α, which was isolated from the venom of the Egyptian Cobra (Naja haje haje), is a small and basic protein consisting of a single polypeptide chain of 61 amino acid residues cross-linked by four disulfide bridges (Botes and Strydom 1969). This protein belongs to the Type I. It has a molecular weight of 6,845.

Pelamis platurus major toxin, a venom of the sea snake belonging to the Hydrophiine subfamily, is a homogeneous and basic proteins with a high isoelectric point, 9.69 (Tu et al. 1975). The venom was extracted from the venom glands of Pelamis platurus snakes captured on the Pacific coast of Costa Rica, Central America. The toxin contains 55 amino acid residues with four disulfide linkages (Tu et al. 1975). The total number of amino acid residues for the Pelamis toxin is a low value when compared to most Type I neurotoxins.

Laticauda semifasciata toxin b, one of toxins of Laticauda semifasciata from the Phillippines which belongs to a subfamily of sea snake, Laticaudinae, has a different amino acid composition from those of the other subfamily. It consists of 61 aminoa cid residues that give it a molecular weight of 6,677 and has a high isoelectric point of 9.3 (Tu et al. 1971).

As compared to non-neurotoxin venoms, neurotoxins possess a neuromuscular blocking property. Neurotoxins specifically bind to the acetyl-choline receptor site at the neuromuscular junction (for further reading, refer to Tu 1973). Utilizing this property many investigators are using snake neurotoxins to elucidate the mechanism of neurotransmission across the synapses or neuromuscular junction (Tu 1973).

The amino acid composition of toxins obtained from the venoms of the Hydrophiinae and Laticaudinae subfamilies are very similar (Tu 1973). Amino acid composition of all sea snake toxins are not only similar among themselves but also to those of Elapidae. The composition of the toxins mentioned above are tabulated in Table 1.

Chemical modification studies of snake venoms suggests that main-backbone conformation and some side-chains are primarily responsible for toxicity of venoms. The disulfide bond is essential for the lethal action of toxins. Yang (1967) reported the lost of toxicity of cobrotoxin from N. naja atra venom by reduction with β-mercaptoethanol or oxidation with performic acid (Yang 1965) of disulfide bonds. Disulfide bonds are primarily responsible for the maintenance of the neurotoxin conformation, thus change in spatial arrangement without altering the primary structure still causes the loss of toxicity (Tu 1973). Tyrosine residue was reported to be essential for maintenance of the toxin

Table 1. Comparison of Amino Acid Composition of Snake Venom Proteins

Genus and Species	*N. naja*	*N. haje haje*	*P. platurus*	*L. semifasciata*
Geographical Origin	India	Africa	Costa Rica	Philippines
Name	Cobramine B	Neurotoxin α	Major Toxin	Toxin b
Lysine	8	6	5	5
Histidine	0	2	2	1
Arginine	2	4	3	5
Aspartic acid	5	7	6	4
Threonine	3	7	7	5
Serine	2	4	5	6
Glutamic acid	0	7	7	8
Proline	4	4	1	4
Glycine	2	5	3	6
Alanine	2	0	1	0
Half-cystine	6	8	8	8
Valine	6	1	1	3
Methionine	2	0	1	0

Table 1. (Continued)

Genus and Species	*N. naja*	*N. haje haje*	*P. platurus*	*L. semifasciata*
Geographical Origin	India	Africa	Costa Rica	Philippines
Name	Cobramine B	Neurotoxin α	Major Toxin	Toxin b
Isoleucine	1	3	2	4
Leucine	5	1	1	1
Tyrosine	3	1	1	1
Phenylalanine	1	0	0	2
Tryptophan	0	1	1	1
Total Residue	52	61	55	61
References	Larsen and Wolff 1968	Botes and Strydom 1969	Tu *et al.* 1975	Tu *et al.* 1971

conformation in cobrotoxin from N. naja atra by Chang et al. (1971). Similar results were obtained for a tyrosine residue of Naja haje (Chicheportiche et al. 1972). Several studies are reporting that tryptophan residue is important to toxicity of venoms (Hong and Tu 1970, Tu and Toom 1971, Tu and Hong 1971, Seto et al. 1970).

The knowledge obtained from model compound studies in Chapters I and II of this paper was used in the conformational analysis of snake venom proteins. Because of the ambiguity in the relationship between backbone conformation of proteins and the amide I and III bands, a predictive computation of protein structure (Chou and Fasman 1974) was performed on neurotoxin α whose primary sequence was determined (Botes and Strydom 1969).

Materials and Methods

The sample of cobramine B was a gift from Dr. J. Wolff of the National Institute of Arthritis, Metabolism, and Digestive Disease. The methods of purification were described by Larsen and Wolff (1968). The solution of cobramine B was prepared by dissolving 20 mg sample in 0.2 ml distilled water and the pH value of the solution was adjusted to 6.95. The solid powder for spectra was obtained by freeze-drying this solution. Glycyl-L-tyrosine was purchased from Sigma Chemical Co. (Cat. No. G3502). These samples were used without further purification.

The sample of neurotoxin α from Egyptian Cobra (Naja haje haje) was a gift from Dr. D. J. Strydom of National Chemical Research Laboratory, Republic of South Africa. The isolation and purification methods were described by Botes and Strydom (1969). The sample preparation for laser Raman spectroscopy is similar to that of cobramine B.

Pelamis platurus major toxin together with Laticauda semifasciata toxin b were obtained from Dr. A. T. Tu of Colorado State University. The purification methods of P. platurus and L. semifasciata are described by Tu et al. (1975, 1971). The solid sample of two toxins are prepared as the same way as cobramine B. Heat denaturation of P. platurus is made by heating a solution of the 5 mg/ml concentration at 100°C for 30 minutes. The heated solution was rapidly cooled in a dry ice-acetone bath and lyophilized immediately. Raman spectra of the lyophilized sample was then taken in the same manner as with the native toxin.

The Raman systems used in the investigation are similar to those described in preceding chapters. The spectra presented in Figures 1, 2, 3 and 4 were recorded by a system equipped with a Spex 1401 double monochromator, ITT FW 130 phototube, Coherent Radiation Model 52G argon-ion laser and standard nuclear counting electronics. The other spectra were obtained on a Spex Ramalog system equipped with a Spex 1401 double monochromator, RCA model C-31034 phototube, Coherent Radiation Model 52G argon-ion laser and Spex

pc-1 photon counting electronics.

All the spectra were obtained with the 514.5 nm line excitation and at about 4 cm^{-1} spectral slit-width resolution.

For experiments above room temperature, the sample cell was sealed and placed in a rectangular slot of a thermostatted brass jacket, the temperature of which was regulated by a Lauda K-2/R circulator.

Predictive computation of protein secondary structure formulated by Chou and Fasman (1974) was performed on neurotoxin α. The computing procedure is well summarized in Chapter IV of this paper.

Results and Discussion

Cobramine B from Indian Cobra (Naja naja) Venom

Spectral Evidence for the Buried Tyrosines in Cobramine B. The spectra of cobramine B in the solid state and in solution are shown in Figure 1 for comparison. The vibrational frequencies in cm^{-1} with assignments are tabulated in Table 1. In Figure 2, the author presents the Raman spectrum of glycyl-L-tyrosine solution at pH 2.50 in the region 600-900 cm^{-1}. This provides a reference spectrum for completely "exposed" tyrosine. In Figure 3, the temperature variation of the intensity ratio of the 644, 830 and 853 cm^{-1} lines in cobramine B is illustrated. Additionally, in Figure 4 the Raman spectra in the tyrosyl region of

insulin, RNase A, and cobramine B were compared to show the dependence of the intensity ratios upon the local environment of tyrosines.

One striking difference between the solid and solution spectra in Figure 1 appears at 644 cm^{-1} due to the in-plane bending vibration of three tyrosines in cobramine B. To understand the nature of this intensity drop upon dissolution, one first needs to know whether the three tyrosines in this protein are "buried" in the interior or "exposed" to the molecular surface. For this purpose, a simple dipeptide, glycyl-L-tyrosine, was employed as a model compound for "exposed" tyrosine. The spectrum in Figure 2 shows that the intensity ratio of the tyrosine lines at 644, 828 and 853 cm^{-1} is roughly 0.7:1.0:1.4. This ratio was found to be quite different from that of cobramine B solution (i.e., 0.5:1.0:0.5) at room temperature (Figure 1(b)). This latter ratio, however, depends on temperature (see Figure 3) and as such provides further evidence for the assignment. At 85°C or above, this ratio approaches that of glycyl-L-tyrosine solution. The author interprets this as due to the thermal denaturation of the protein. At this temperature, the author believes, the "buried" tyrosines became "exposed" to the solvent. It should be noted that an intensity increase of tyrosine lines at 836 and 856 cm^{-1} was also observed by Brunner and Sussner (1972) when lysozyme was thermally denatured. Based on the above evidence, it

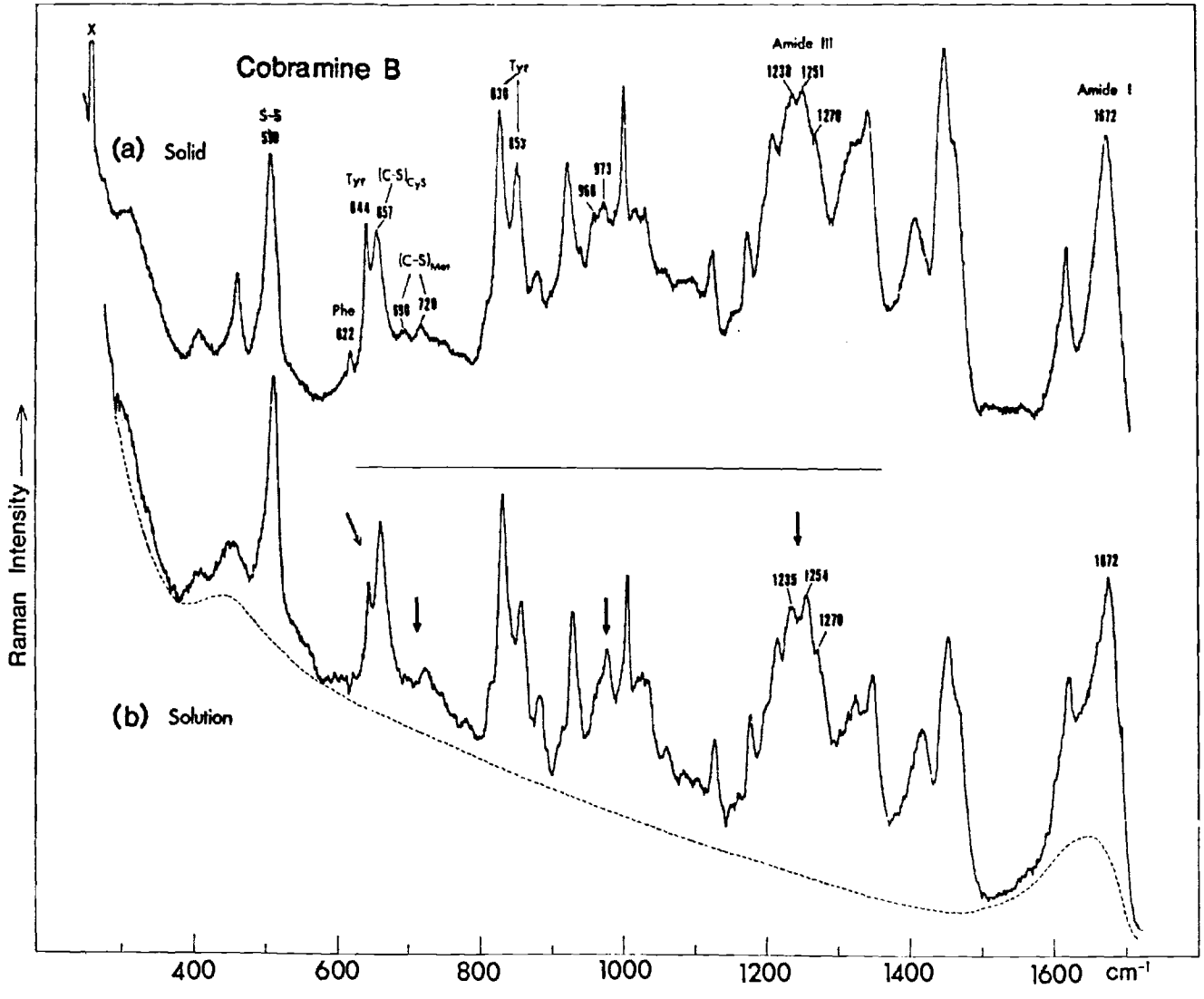

Figure 1. Raman Spectra of Cobramine B in Solid and in Solution

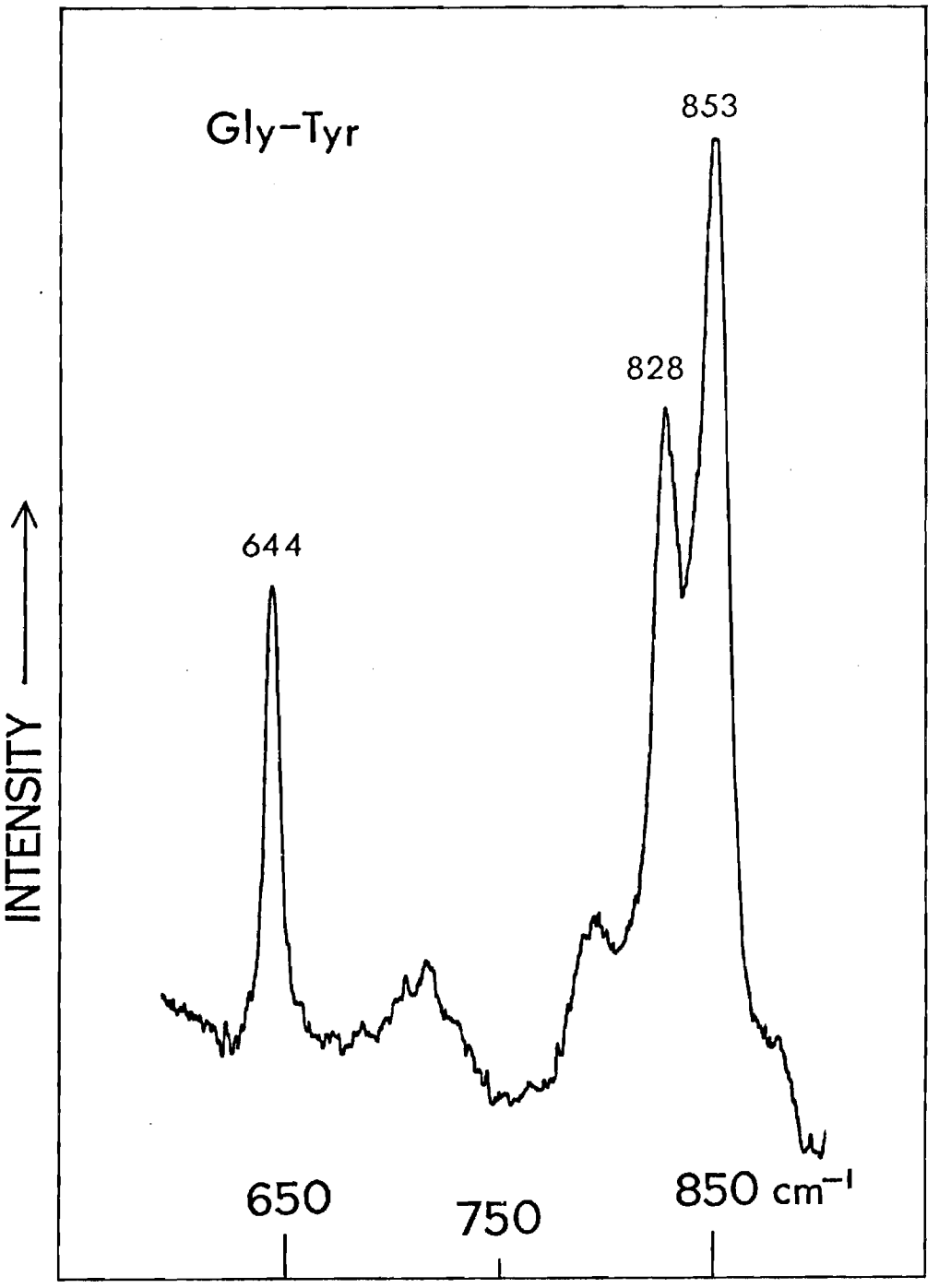

Figure 2. Raman Spectrum of Glycyl-L-Tyrosine Solution (600-900 cm^{-1})

appears that some tyrosines in cobramine B are probably "buried".

There is a question as to how many "buried" tyrosines there are in the native state. Yu and Jo (1973) have looked at the tyrosine ratio in neurotoxin α, a protein containing a single tyrosine. If the tyrosine is "exposed", it would be expected that the ratio is 0.7:1.0:1.4 and if it is "buried", a ratio closer to that of cobramine B, 0.5:1.0:0.5. Our observations showed an intensity ratio of 0.5:1.0:0.5 in this single tyrosine, indicating that this may be a standard ratio for "buried" tyrosines in proteins. On this basis the author would argue that all three tyrosines in cobramine B are "buried".

Thus, the intensity ratio of these three lines may provide a reliable index of the environments and interactions of tyrosines in a protein. When tyrosines are completely "exposed" to water, one may expect a ratio of 0.7:1.0:1.4 as in glycyl-L-tyrosine solution. However, the "buried" tyrosines in a protein may not always give rise to the intensity ratio of exactly 0.5:1.0:0.5. The altered intensity ratio may be the manifestation of altered specific interactions involving tyrosines such as those due to the difference in conformation between phases.

The conformation of proteins in lyophilized state may be different from that in solution. This has been demonstrated from Raman spectra of RNase A and lysozyme. Also

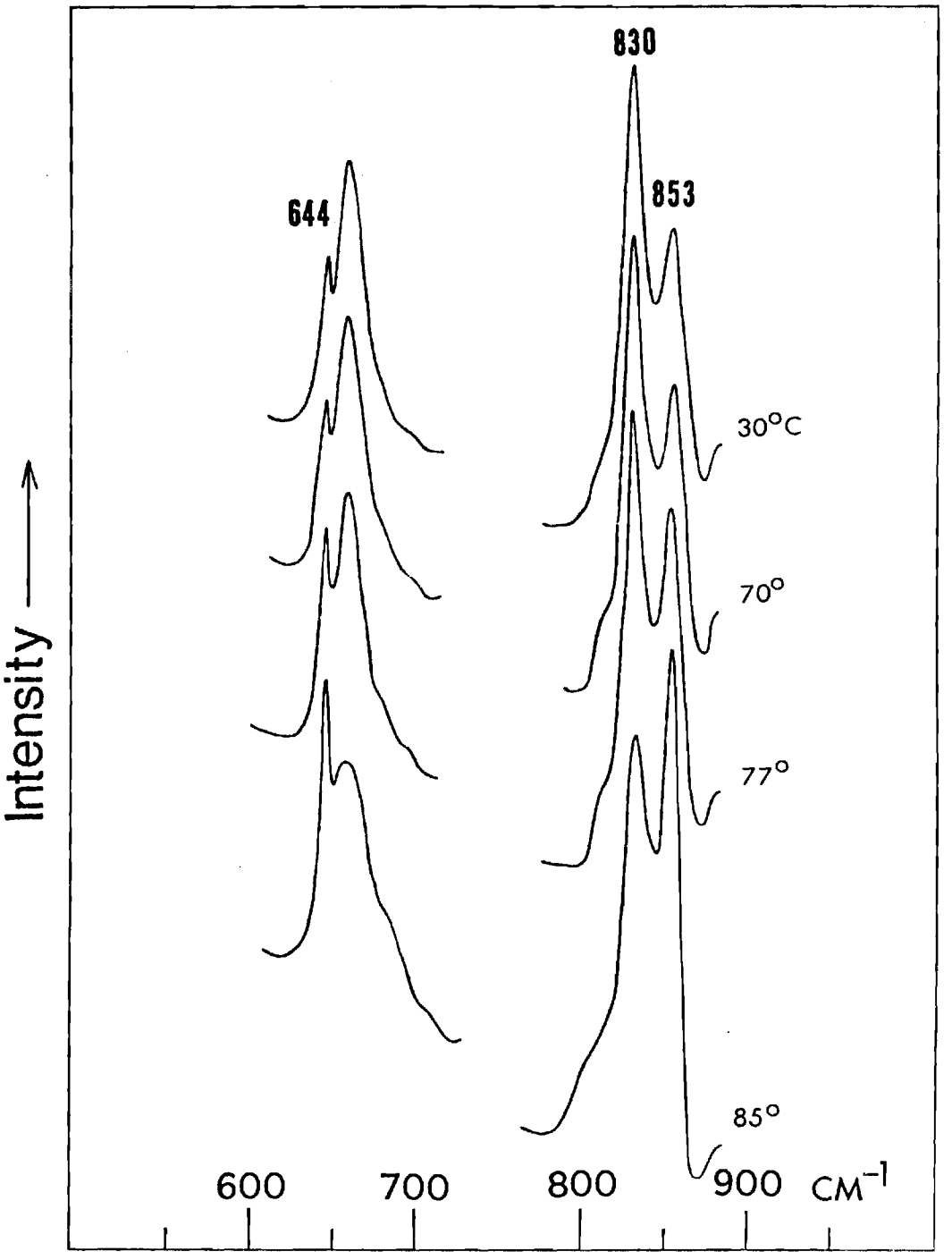

Figure 3. Temperature Dependence of Raman Scattering Intensities of Tyrosyl Lines of Cobramine B

the studies on crystalline RNase A obtained by ethanol precipitation indicate that the spectra of crystalline and aqueous RNase A are different in the tyrosyl ring vibration region.

Here, several spectral differences are observed between the solid and aqueous spectra (Figure 1). The relative scattering intensities of the tyrosyl lines of lyophilized cobramine B do not fit the 0.5:1.0:0.5 ratio although it is concluded that all tyrosines in this protein are probably "buried". The intensity increase of the 644 cm^{-1} line with solid over that in the solution may be due to the conformation-induced changes in the local environment or interaction of tyrosine(s). One possible reason for such changes might be the weakening or breaking of hydrogen bonds involving the phenolichydroxyl group.

A similar intensity increase of 644 cm^{-1} line upon solidification of insulin and RNase A is shown in Figure 4. In the case of RNase A, the line at 852 cm^{-1} has also increased its intensity on solidification. The tyrosine ratio of aqueous RNase A is approximately 0.6:1.0:0.8, a ratio one would expect on the basis of three "buried" and three "exposed" tyrosines. In insulin solution and crystals, the tyrosine ratio is close to that of glycyl-L-tyrosine solution, suggesting that tyrosines may not be completely "buried". In fact, the X-ray diffraction studies of Blundell and co-workers (1972a, b) indicate that none of the four tyrosines

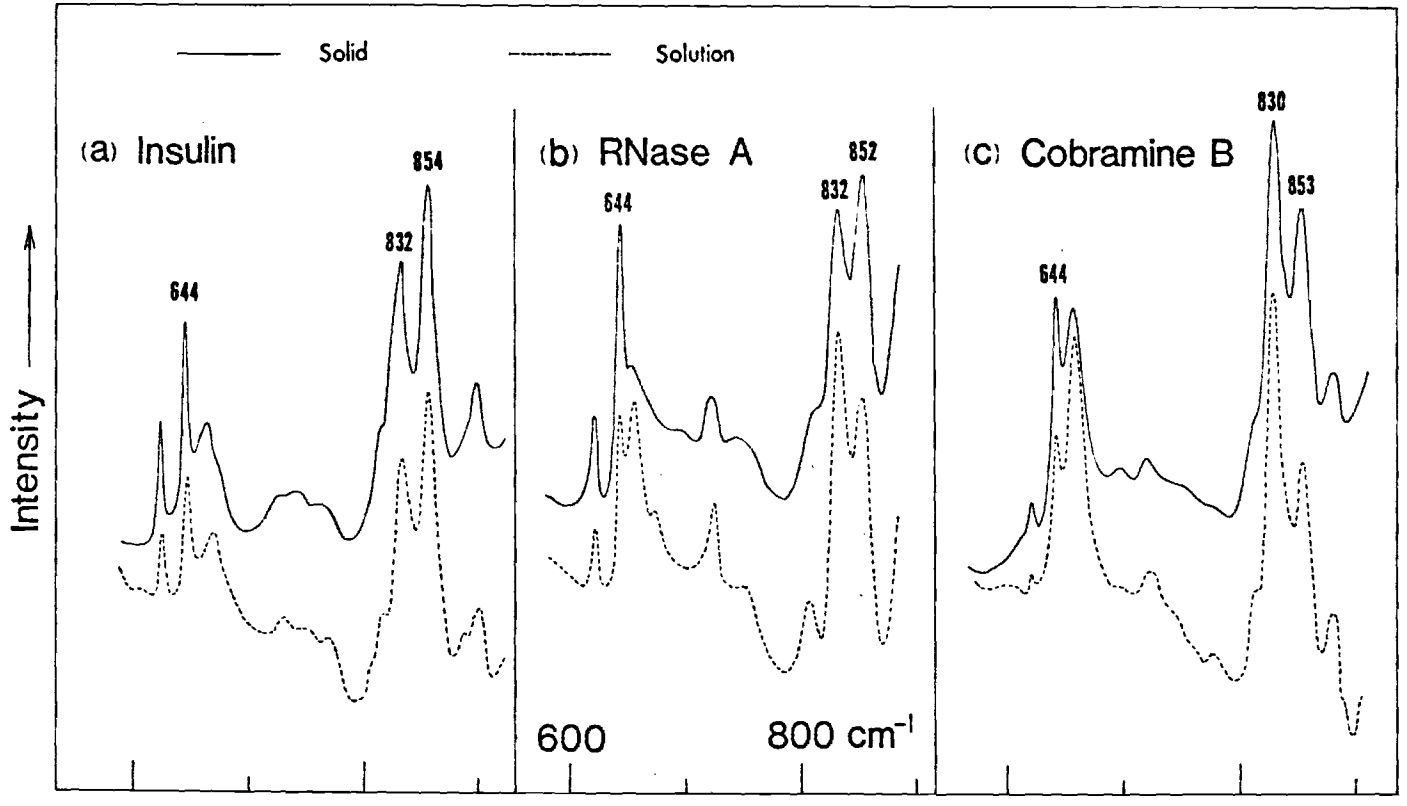

Figure 4. Comparison of Raman Spectra of Insulin, RNase A and Cobramine B in Tyrosyl Ring Vibrations (600-900 cm^{-1})

Table 2. Raman Spectra of Cobramine B

Frequencies (cm^{-1})		Assignments
Solid	Solution	
315 (1.5B)	310 (1.2)	
410 (1.1B)	410 (1.1)	
463 (2.3)	455 (1.0B)	
510 (6.2)	511 (6.9)	ν(S-S)
622 (0.8)	622 (0.3)	Phe
644 (4.1S)	644 (2.6S)	Tyr
657 (0.5)	657 (4.3)	ν(C-S) Cys
698 (0.5)		ν(C-S) Cys or Met
720 (0.8)	725 (1.3)	
	779 (0.4)	
810 (1.3)	810 (1.5)	
830 (6.7)	831 (7.0)	Tyr
853 (3.6)	856 (3.9)	
881 (1.0)	883 (1.5)	
	913 (1.0sh)	
923 (3.6)	929 (4.3)	ν(C-C)
960 (1.0)	963 (1.0)	
973 (1.3)	975 (2.0)	
1003 (3.7S)	1004 (3.4S)	Phe
1018 (1.0)	1018 (1.0)	

Table 2. (Continued)

Frequencies (cm^{-1})		Assignments
Solid	Solution	
1031 (1.5)	1031 (1.5)	Phe
1058 (0.7)	1060 (0.9)	
1095 (0.8)	1084 (0.8)	ν(C-N)
	1103 (0.4)	
1125 (2.3)	1125 (2.3)	
1174 (2.0)	1176 (2.1)	Tyr
1195 (1.0 sh)	1196 (1.0 sh)	Phe
1210 (3.6)	1212 (3.3)	Tyr
1238 (4.6)	1235 (4.6)	Amide III
1251 (4.6)	1254 (4.9)	
1270 (2.3)	1270 (2.3)	
1320 (3.1)	1321 (2.6)	-CH deformation
1341 (5.2)	1345 (4.1)	
1407 (3.1)	1413 (3.4)	symmetrical $-CO_2$ stretching
1448 (10.0)	1450 (7.4)	$-CH_2$ and $-CH_3$ deformation
1460 (6.2 sh)	1463 (5.2 sh)	
1603 (2.1 sh)	1603 (2.1 sh)	Phe and Tyr
1615 (4.6)	1615 (4.6)	Tyr
1672 (8.9)	1672 (8.9)	Amide I

Table 2. (Continued)

Vibrational frequencies are expressed as the displacement in cm^{-1} of the Raman lines. Abbreviations: B = broad, S = sharp, sh = shoulder, ν(X-Y) = stretching vibration of X-Y. Numerical figures in parentheses represent relative intensity based on a value of 10.0 for the strongest line at 1448 cm^{-1}.

in insulin crystals are completely buried.

In Figure 3, the temperature dependence of the Raman line intensity of three tyrosyl rings at 644, 830 and 853 cm^{-1} is demonstrated. Up to 85°C, the relative intensities of these lines are continuing their changes to the ratio of that of glycyl-L-tyrosine solution. This fact implies that three tyrosine side-chains are exposed to the surface of the molecule.

Conformational Information of S-S and C-S Stretching Vibrations. According to Sugeta et al. (1972, 1973), the line at 510 cm^{-1} is expected for the disulfide bond in the gauch-gauch-gauch form. The line at 510 cm^{-1} in the solid and solution spectra can be assigned to the gauch-gauch-gauch form of three disulfide bonds. The relatively sharp and symmetrical features indicate that the conformational geometry about the three S-S bonds is uniform in the solid and solution states. For the region of 650 to 750 cm^{-1}, the C-S stretching vibrations of the half-cystine and methionine side-chains are responsible. The line at 698 cm^{-1} may be tentatively assigned to P_N form of any half-cystine group(s) or to methionine group(s) in the trans-gauch form. From studies of model compounds, Nogami et al. (1975) showed that the methionine of the trans-gauch form should have another line around 750 cm^{-1}. In the solid spectrum (Figure 1a), a small bump is observed at 750 cm^{-1}. The author assigns the lines of 698 and 750 cm^{-1} to the C-S

stretching vibrations of the methionine group(s) in the trans-gauch form. But a possibility cannot be denied that there may exist a half-cystine group in the P_N form which could be also responsible to the 698 cm^{-1} line. The line at 657 cm^{-1} in the solid spectrum seems largely due to the P_H form of half-cystine side-chain. However, any methionine group in the gauch-gauch form also should be responsible for the line of 657 cm^{-1}. Because it is expected that the lines from the gauch-gauch form of methionine group appears at 645 and 723 cm^{-1} frequencies (Nogami et al. 1975). It is also believed that the 720 cm^{-1} line may be contributed by any P_C form of half-cystine group. Thus, the author tentatively assigns the line 720 cm^{-1} to the C-S stretching vibration of the methionine side-chain(s) in the gauch-gauch form or/and possibly of half-cystine groups in the P_C form. Therefore, for the protein which contains both methionine residues and half-cystine groups it is hard to make assignments of C-S stretching bands. However, from the fact that snake venoms which do not contain any methionine groups also show a peak at 657 cm^{-1} with a common feature as shown in the later sections of this Chapter, it is suggested that the 657 cm^{-1} band may be only due to the half-cystine groups in P_H. Then two methionine side-chains may be each in gauch-gauch and trans-gauch forms. In the solution spectrum, the line at 698 and 750 cm^{-1} has become barely visible with emerging more intense peaks at 657 and 720 cm^{-1}. This

phenomena may suggest that the line at 698 cm^{-1} in the solid spectrum was contributed by the C-S stretching vibration of the methionine group in the trans-gauch form and 720 cm^{-1} by the methionine group in the gauch-gauch form and upon dissolution, changes in the local environment of one methionine group is opposite to the other. It can not be denied that the half-cystine groups in P_H (657 cm^{-1}) may have changes in their local environment. The fact that the local environment of half-cystine or methionine side-chains is sensitively reflected in the intensity variation of Raman bands of C-S stretching vibration was mentioned in the RNase A case of Chapter II of this dissertation. From the amide III and I and S-S stretching vibration regions of the solid and solution spectra, it is hard to believe that a transformation of conformation about half-cystine groups occurrs without any marked changes of backbone conformation and cystine linkage of the protein.

Recently, Brunner and Sussner (1972) observed that intensity of the S-S stretching vibration at 504 cm^{-1} in the Raman s pectra of lysozyme solution decreased on heating and was essentially zero above the denaturation temperature of 76°C. On the basis of the evidence, they concluded that all disulfide linkages have been broken in the denatured form. On the contrary, Chen _et al._ (1973) observed that all the disulfide bonds are intact and their geometry is little changed up to 76°C. To determine whether the disulfide bonds

of cobramine B might also be affected at its denaturation temperature of about 85°C, the author has investigated the intensity of the S-S line at 510 cm^{-1} as a function of temperature (30-85°C). In the case of cobramine B, the S-S line intensity has not decreased appreciably up to 85°C, indicating that all disulfide bonds still remain intact when this protein is thermally denatured.

The Information about Backbone Conformation. Yu and Liu (1972), based on their Raman studies of glucagon in various conformational states, concluded that the α-helical, random-coiled (H-bonded), and antiparallel-β structure of a protein should have the amide I frequencies at 1660, 1665 and 1672 cm^{-1}, respectively and that the corresponding amide III vibrations should be 1266, 1248 and 1232 cm^{-1}. On the other hand, Chen et al. (1973, 1974) assigned the 1672 cm^{-1} line to random-coil conformation in their studies of lysozyme denaturation. But the argument is still not persuasive. Therefore, the author rather relies on the amide III region in this case. In the present spectra of Figure 1, three resolved peaks are observed at 1238, 1251 and 1270 cm^{-1} in solid and at 1235, 1254 and 1270 cm^{-1} in solution indicating the coexistence of β structure, random-coil and α-helices. Although the amide I band appears at 1672 cm^{-1}, the shape is broader than that observed by Yu and Liu (1972). It may be due to a large portion of random-coil structures as reflected in the amide III region. Since the small shifts in the

amide III peaks between the solid and solution spectra are within the uncertainty of ±4 cm^{-1}, it is not clear that the shifts may be due to subtle changes in the polypeptide backbone.

Also, in the C-C-N skeletal stretching region, the doublet at 960 and 973 cm^{-1} changes in the relative intensity. This may be the reflection of conformational change.

Neurotoxin α from Egyptian Cobra (Naja haje haje) Venom

This toxin has one mole of tyrosine as shown in the amino acid analysis (Table 1). The laser Raman spectra has revealed the presence of tyrosine as peaks at 644, 834 and 846 cm^{-1}. The relative intensity ratio of 644, 834 and 846 cm^{-1} in solid and solution are not much different from that of cobramine B. Thus, it is concluded that the single tyrosine residue is involved in some intramolecular interaction and is not readily accessible to water molecules.

Most of the snake venom neurotoxins, either Type I or Type II, contain one or two residues of tryptophan, one residue being predominant (Tu 1973). It is well documented that the single tryptophan residue is essential for toxic action as selective modification of tryptophans causes a loss of toxicity (Hong and Tu 1970, Tu and Toom 1971, Tu and Hong 1971, Seto et al. 1970).

The 1361 cm^{-1} band of tryptophan in laser Raman spectra is very sensitive to the tryptophan being "buried"

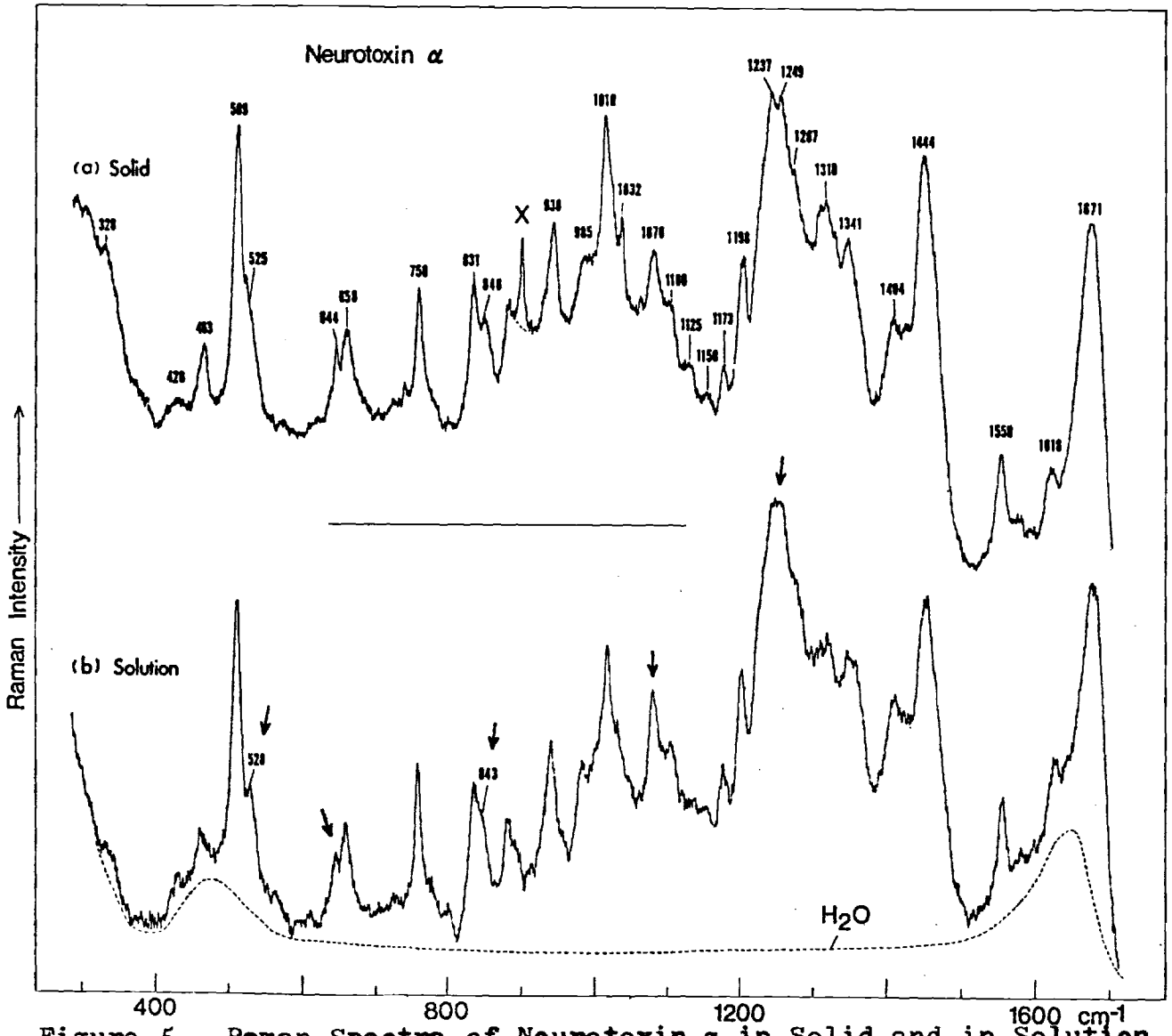

Figure 5. Raman Spectra of Neurotoxin α in Solid and in Solution

Table 3. Raman Spectra of Neurotoxin α

Frequencies (cm^{-1})		Assignments
Solid	Solution	
328(0.6)	328 (0.6B)	
428 (0.8B)	428 (0.8)	
463 (3.1)	463 (1.6B)	
509 (10)	509 (10.6)	ν (S-S)
525 (4.7 sh)	528 (4.4)	
644 (3.1)	644 (3.12)	Tyr
658 (3.4)	658 (3.8)	ν (C-S) Cys
701† (0.5)		
724† (0.8)	726† (0.9)	
746† (1.1)		
758 (5.0)	758 (5.6)	
831 (5.0)	831 (4.7)	Tyr
846 (3.8)	843 (4.4)	
881 (2.2)	881 (3.4)	Trp
938 (6.9)	938 (6.6)	ν(C-C)
985 (3.8)	985 (4.7)	
1010 (6.9)	1010 (6.6)	Trp
1032 (1.7)	1032 (1.4)	
1076 (3.1)	1076 (3.8)	ν(C-N)
1100 (1.7)	1100 (1.7)	
1125 (1.0)		

Table 3. (Continued)

Frequencies (cm^{-1})		Assignments
Solid	Solution	
1150 (0.6)	1150 (0.6)	
1173 (1.6)	1173 (2.2)	Tyr
1198 (5.3)	1198 (5.3)	Tyr
1237 (11.3)		
1249 (10.9)	1243 (11.5)	Amide III
1267 (6.9)	1267 (6.9)	
1310 (4.4)	1310 (4.4)	-CH deformation
1341 (4.4)	1341 (4.4)	
1404 (4.7)	1404 (5.6)	symmetrical-CO_2 stretching
1444 (10)	1444 (7.5)	-CH_2 and -CH_3 deformation
1550 (3.8)	1550 (3.1)	
1618 (2.5)	1618 (2.5)	Tyr
1671 (10.9)	1671 (10.9)	Amide I

†: frequency uncertain
For other notations, see Table 2 of this Chapter.

or "exposed" (Chen et al. 1973, Yu 1974). When the indole ring is buried and involved in certain interactions inside the core of a protein molecule, the 1361 cm^{-1} band shows a sharp peak. As the indole ring becomes accessible to the solvent molecule, the 1361 cm^{-1} line diminishes. Random-coiled glucagon does not show a sharp line at 1361 cm^{-1}. The lack of a distinct peak at 1361 cm^{-1} in the spectra of Figure 5 suggests that the single tryptophan residue is exposed. The results of the laser Raman study agree well with the results of chemical modification performed by many investigators (Yu 1973, Tu 1974, Tu and Hong 1971, Tu and Toom 1971, Tu et al. 1971). Amino acid analysis of the neurotoxin α showed no evidence of phenylalanine (Table 1). Laser Raman spectra in Figure 5 also indicates the absence of phenylalanine in the toxin.

There are four disulfide linkages in Type I neurotoxins. It has been observed that the peaks due to the disulfide stretching vibration mode is sensitive to the geometry of the cystine groups. The lines at 509 and 525 cm^{-1} in the solid and solution spectra suggest that the disulfide bonds in the toxin consist of two different type of geometry, i.e., gauch-gauch-gauch and gauch-gauch-trans forms. Furthermore, the narrow and sharp line at 509 cm^{-1} may reflect that the geometry about disulfide bonds belonging to the gauch-gauch-gauch form is in a rather uniform state. Although quantitative assessment of each form is hard to be

made, it may be possible to provide an approximate estimation by examining the relative intensity of these lines of a spectrum of the model compound, lysozyme, shown in Chapter I—one is in the gauch-gauch-trans form and the rest of them in the gauch-gauch-gauch form.

The toxin does not contain any methionine residues as shown in the Table 1. This fact provides advantage in analyzing the Raman band of the C-S stretching vibration of half-cystine side-chain, because the C-S stretching vibration bands of the cystine groups are often overlapped by that of the methionine side-chains. As mentioned earlier, the line, 658 cm^{-1} in the solid and solution spectra can be assigned to the C-S stretching vibration mode of the half-cystine groups in the P_H form. The weak peaks at the 700-730 cm^{-1} region in solid spectrum are at the noise level and do not deserve discussion. In solution, the peak of 658 cm^{-1} increased a little. This may indicate a subtle change in the local environments of the C-S bonds in the P_H form. But it may be reasonable to state that most of the half-cystine groups in this toxin are in the P_H form although the spectrum (700-750 cm^{-1}) shows a sign that a small fraction of other forms might exist.

There appear three peaks in the amide III region, doubly splitted bands at 1237 and 1249 cm^{-1} and one at 1267 cm^{-1}. It has been demonstrated from laser Raman studies of other proteins, such as RNase A and lysozyme, that an α-helix

would appear at a higher wave number, usually above 1260 cm^{-1}. Therefore, the line at 1267 cm^{-1} may be assigned to the α-helix fractions of the peptide backbone. The main lines at 1237 and 1249 cm^{-1} can be assigned to the anti-parallel-β and random-coil structures. The amide I band is shown at 1671 cm^{-1} as a less sharp peak than that of the antiparallel-β structure of glucagon (Yu and Liu 1972). The spectral features of the amide III and I regions are very similar to the case of cobramine B indicating a large fraction of random-coil conformation. Recently, Yu et al. (1975) have reported Raman data of neurotoxins of sea snakes, L. hardwickii and E. schistosa and observed similar features at the amide III and amide I regions as that of neurotoxin α. They concluded that those toxins have large portion of antiparallel-β structure in their backbone conformation. It seems too early to make such conclusion before additional structural data from other physical methods are available. The amide III region of the solution spectrum shows a subtle change in band shape from that of the solid. But the amide I peak still remains the same in both phases. This fact implies that solvation does not have much effect on the backbone conformation of the protein.

Recently, Chou and Fasman (1974) has established an empirical method in predicting secondary structures of proteins whose primary sequences are known. They claim that their accuracy in predicting conformation is better than

80% with respect to X-ray diffraction results. Recently, this method has been demonstrated to give better accuracy than several other prediction methods (Schulz et al. 1974). The method is primarily based on the statistics. The author applied this calculation technique to the neurotoxin α since its chemical sequence is known (Botes and Strydom 1968).

The analyzed results obtained by the above method for neurotoxin α are summarized in Table 4. As seen in Figure 6, the only region with a clustering of more than four helical residues is 1-7 where there are five residues out of seven residues with an average helical potential of $\langle P_\alpha \rangle = 1.08$ (condition A-1). This region also has high potential for β sheet nucleation since five out of seven are β residues having $\langle P_\beta \rangle = 1.08$ (condition B-1). However, this region with $(Hh_4ib)_\alpha$ and $(h_5b_2)_\beta$ assignments is predicted to be α since there are one strong α former (H_α) and one breaker (b_α) compared to no strong β former (H_β) and two β breakers (b_β). This α-helix is terminated by the tetrapeptide of 8-11 whose $\langle P_\alpha \rangle = 0.84 < 1.00$ (condition A-2). β regions are located at 13-17 and 51-56 with $\langle P_\beta \rangle = 1.12$ and 1.21 respectively. Although the residues, 49-56, have a value of $\langle P_\beta \rangle = 1.17$ which is greater than 1.05, including the Glu 50 in the region violates the condition B-3. The region of 21-25 having $(h_3b_2)_\beta$ and $\langle P_\beta \rangle = 1.04$ satisfies the condition B-1 but violates the rule 2 requiring that $\langle P_\beta \rangle \geq 1.05$. Therefore, the neurotoxin α contains 11% α-helix, 18% β

Neurotoxin α

	Loop 1			Loop 2							
	1	2	3	4	5	6	7	8	9	10	11
H_2N -	Leu	Gln	Cys	His	Asn	Gln	Gln	Ser	Ser	Gln	Pro
α	[H	h	i	h	b	h	h]	i	i	h	B
β	h	h	h	b	b	h	h	b	b	h	b

						Loop 3				
	12	13	14	15	16	17	18	19	20	21	22
	Pro	Thr	Thr	Lys	Thr	Cys	Pro	Gly	Glu	Thr	Asn
α	B	i	i	I	i	i	B	B	H	i	b
β	b	[h	h	b	h	h]	b	i	B	h	b

	Loop 4									
	23	24	25	26	27	28	29	30	31	32	33
	Cys	Tyr	Lys	Lys	Arg	Trp	Arg	Asp	His	Arg	Gly
α	i	b	I	I	i	h	i	i	H	i	B
β	h	h	b	b	i	h	i	i	b	i	i

Figure 6. Predictive Anaylsis of Neurotoxin α Secondary Structures

	Loop 5										
	34 Ser	35 Ile	36 Thr	37 Glu	38 Arg	39 Gly	40 Cys	41 Gly	42 Cys	43 Pro	44 Ser
α	i	I	i	H	i	B	i	B	i	B	i
β	b	H	h	B	i	i	h	i	h	b	b

	Loop 6										
	45 Val	46 Lys	47 Lys	48 Gly	49 Ile	50 Glu	51 Ile	52 Asn	53 Cys	54 Cys	55 Thr
α	h	I	I	B	I	H	I	b	i	i	i
β	H	b	b	i	H	B	[H	b	h	h	h

	Loop 7				Loop 8		
	56 Thr	57 Asp	58 Lys	59 Cys	60 Asn	61 Asn	- OH
α	i	i	I	i	b	b	
β	h]	i	b	h	b	b	

Figure 6. (Continued)

Table 4. Conformational Prediction for Neurotoxin α

Residues	Predicted	$<P_\alpha>$	$<P_\beta>$
1 - 7	α	1.08	1.08
8 - 12	C	0.78	0.78
13 - 17	β	0.86	1.12
18 - 50	C	0.89	0.90
51 - 56	β	0.81	1.21
57 - 61	C	0.86	0.82

Predicted β-turn	$p_t \times 10^4$	$<P_t>$	$<P_\alpha>$	$<P_\beta>$
11 - 14	0.766	1.27	0.70	0.91
17 - 20	0.803	1.21	0.86	0.75
21 - 24	0.604	1.28	0.73	1.11
24 - 27	0.736	1.07	0.89	0.92
31 - 34	1.23	1.73	0.84	0.79
38 - 41	0.576	1.38	0.66	0.96
40 - 43	0.551	1.39	0.67	1.01
42 - 45	0.667	1.14	0.82	1.07
57 - 60	0.790	1.28	0.89	0.87

For notations, see Table 1 of Chapter IV of this dissertation.

sheets and 71% random-coil structures in the prediction. Computation was extended to β-turn prediction. The relative probability that a tetrapeptide will form a β-turn was computed for 58 combinations in 61 amino acid residues of the neurotoxin α protein. Among them, the tetrapeptides having the value of p_t greater than 0.5×10^{-4} are listed in Table 4. $p_t = 0.5 \times 10^{-4}$ was chosen as a reasonable cut-off value in predicting β-turn (Chou and Fasman 1974). The tetrapeptide of 51-54 has a p_t value (0.578×10^{-4}) greater than the cut-off value but was eliminated since the region was already counted as a β region. A number of β-turns and high portion of random-coil structures in the predictive conformation may be in good agreement with two dimensional structure described as having eight loops by Tu (1973) if the shape of loops is regarded as round type rather than hair-pin (β structure). The schematic diagram of the secondary structure predicted in neurotoxin α is presented in Figure 7.

Based on the above results and experimental date (Yu and Liu 1972, Yu and Jo 1973a, b), the amide III bands at 1237, 1249 and 1267 cm^{-1} in Figure 5 could be assigned to β, random-coil and α structures respectively. The fact that the 1237 cm^{-1} line assigned to β structure appears as intense as random-coil band at 1249 cm^{-1} in Figure 5 although the portion of β sheet is small relative to random-coil structure may be the result of peak overlapping of a weak β band

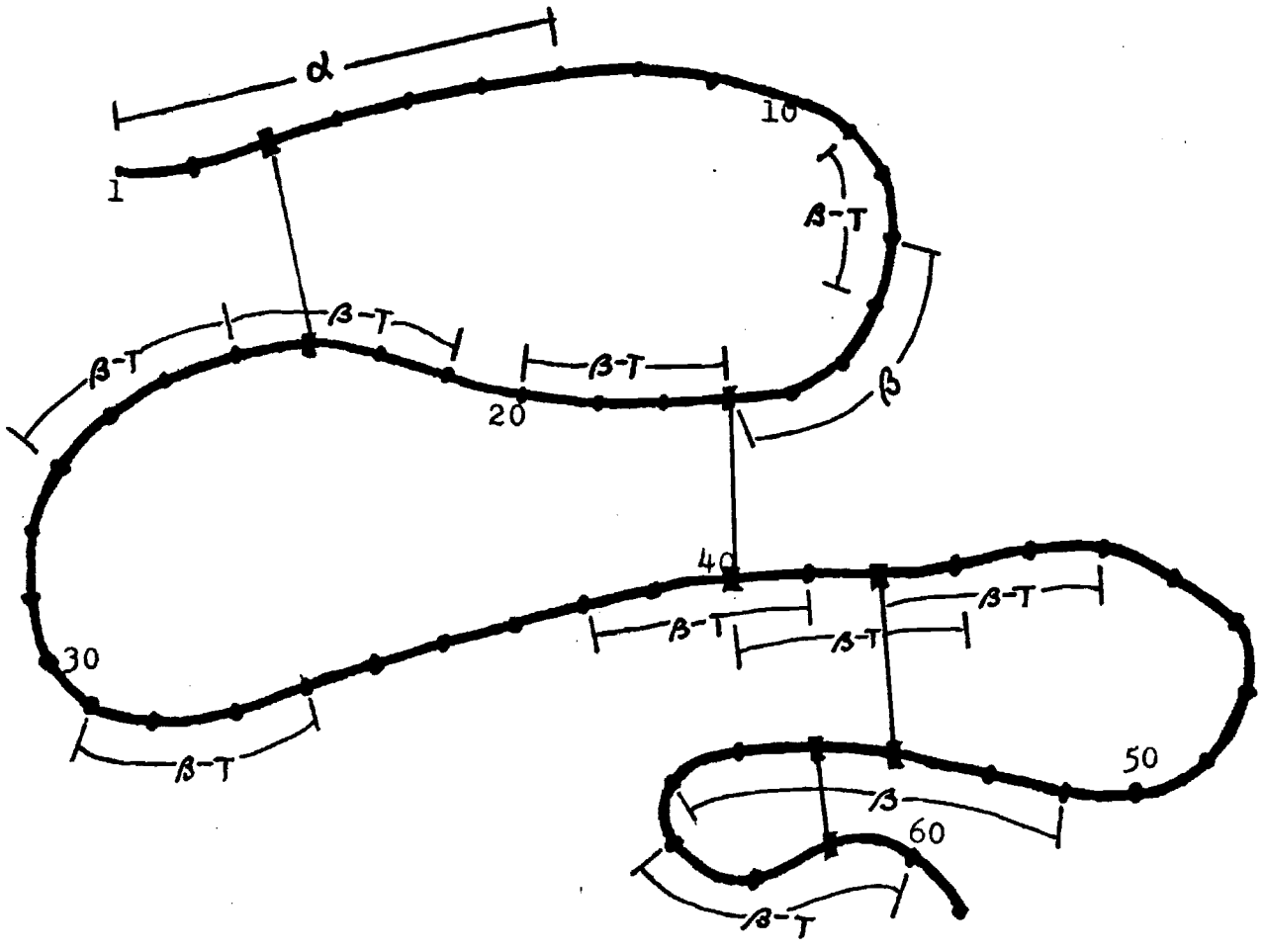

Figure 7. Schematic Diagram of Neurotoxin α Backbone

with an intense and broad random-coil peak.

Pelamis platurus Major Toxin and Laticauda semifasciata Toxin b

The conformation of isolated Pelamis major toxin (toxin a) was studied by laser Raman spectroscopy. It is evident from the spectrum of native Pelamis major toxin (Figure 8) that the peptide backbone is similar to that of cobramine B or neurotoxin α and possibly contains a large fraction of random-coil. This conclusion is based on the distinctive amide I band that appears at 1672 cm^{-1} and the amide III band that appears at 1245 cm^{-1}. This observation is essentially identical to those of other sea snake toxins such as Lapemis hardwickii and Enhydrina schistosa (Yu et al. 1975). This is rather reasonable as all three sea snakes, P. platurus, L. hardwickii and E. schistosa all belong to the same subfamily of Hydrophiinae. As mentioned earlier, Laticauda semifasciata belongs to a different sea snake subfamily, Laticaudinae, and the amino acid compositions of its toxins are different from those of the other subfamily. Comparison of the Raman spectra of toxins from these two subfamilies provides new insight into the structure of the pure toxins.

The amide III and I bands of the L. semifasciata toxin spectrum is actually same as that of the P. platurus toxin but shows a little broadening in shape and frequency shift to 1248 cm^{-1} of the amide III region indicating a

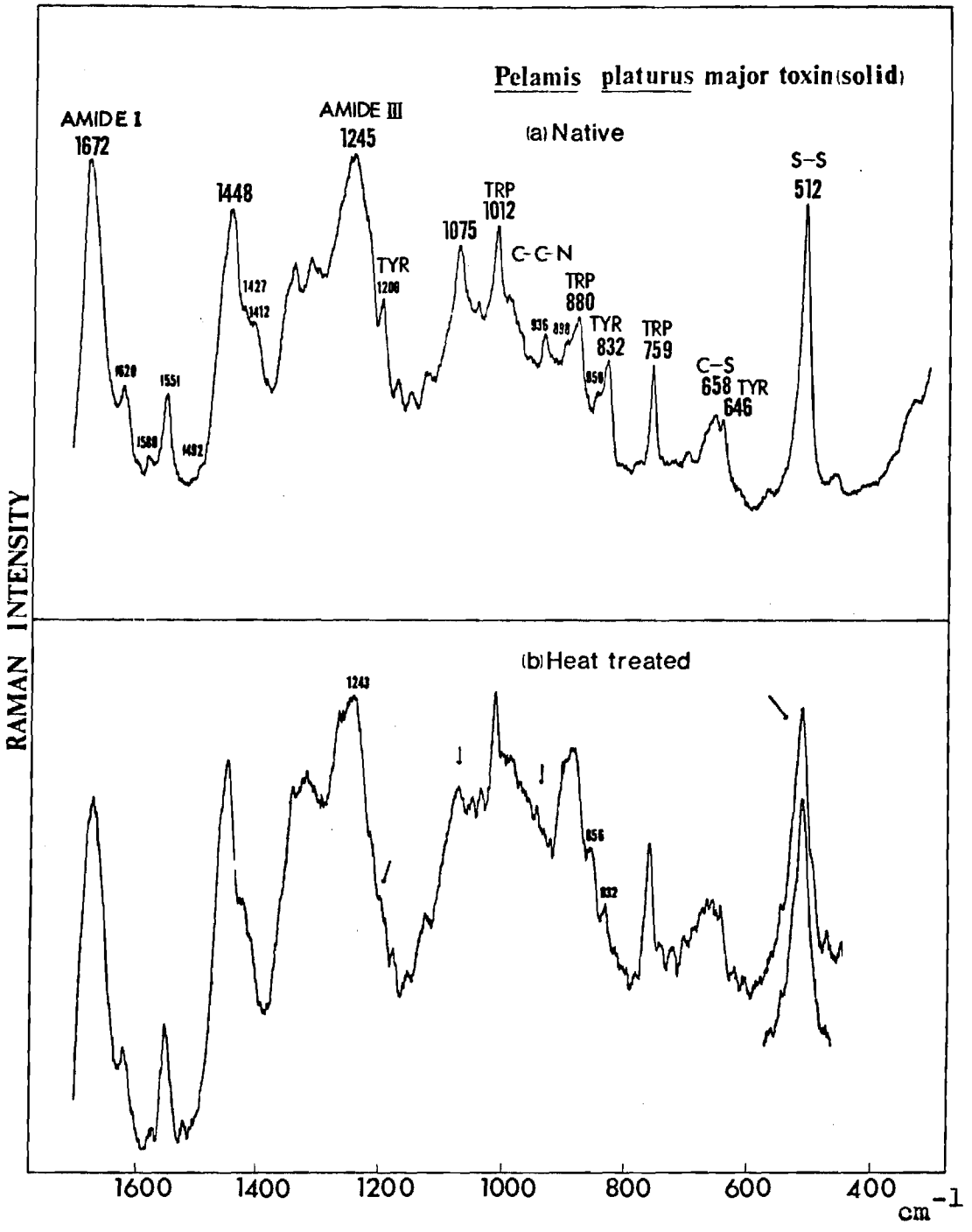

Figure 8. Raman Spectra of P. platurus Major Toxin and Its Heat-treated Powders

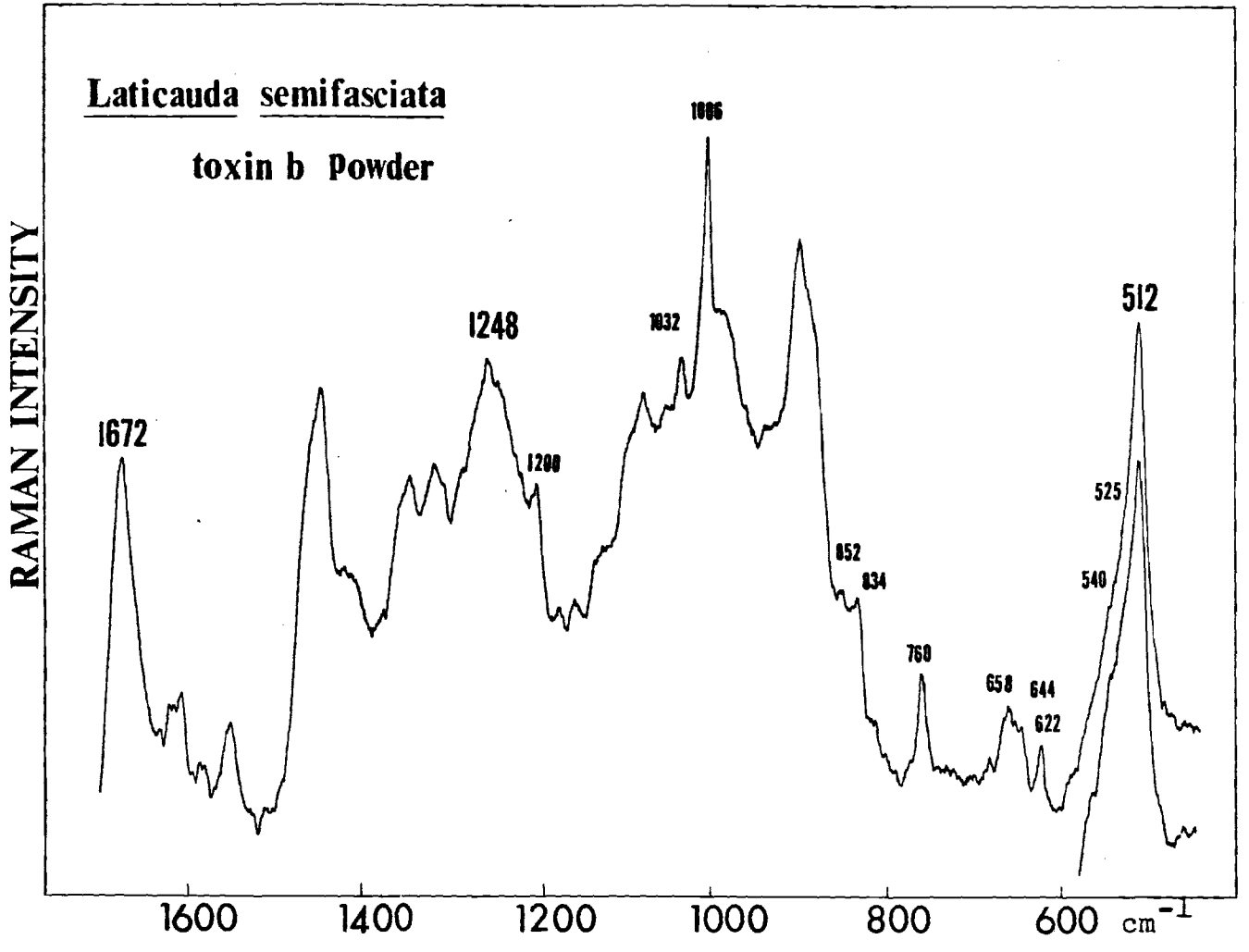

Figure 9. Raman Spectrum of L. Semifasciata Toxin b Powder

large fraction of random-coil backbone conformation having hydrogen bonds in less uniform strength.

The study of glycyl-tyrosine and lysozyme (Yu and Jo 1973a, Chapter I of this dissertation) showed that the relative intensities of the Raman lines at 846 and 836 cm^{-1} are related to the environment of the tyrosine side-chain. Figures 8(a) and 9 show that the Raman spectra of P. platurus toxin and L. semifasciata toxin have differences in the intensity of the peaks at 852 and 835 cm^{-1}. In contrast to the Raman spectra of toxins from P. platurus, L. hardwickii and E. schistosa (Yu et al. 1975), the spectrum of toxin b from L. semifasciata indicates the intensity of the 852 cm^{-1} line is somewhat greater than that of the 834 cm^{-1} line. This indicates that the environments of tyrosine side-chains in this toxin is somewhat different from that of those former toxins. Judging from the intensity ratio of tyrosine bands at 850 and 832 cm^{-1} in the solid spectrum of native Pelamis toxin, it is apparent that the single tyrosine is not readily accessible to the surface of the molecule and perhaps is involved in some kind of binding with other side-chain residues. The results of the laser Raman spectra of L. hardwickii (Yu et al. 1975) are in good agreement with the fact that only 50% of the tyrosine molecule is nitrated with tetranitromethane (Raymond and Tu 1972). The opposite feature in the spectrum of Laticauda toxin suggests that the specific interaction of the tyrosine aromatic ring with its

neighboring group is weak or does not exist and the tyrosyl ring is exposed.

The lack of distinct peak at 1361 cm^{-1} demonstrates that the single tryptophan present in the toxins isolated from P. platurus and L. semifasciata is relatively exposed. For the case of L. hardwickii and E. schistosa toxins, the same observation was made.

A sharp and symmetrical peak at 512 cm^{-1} in the spectrum of P. platurus toxin indicates that the four disulfide bonds present in native toxin have very similar geometry, i.e., the gauch-gauch-gauch form. On the other hand, the spectrum of L. semifasciata toxin b shows shoulders at 525 and 540 cm^{-1} besides the line of 512 cm^{-1}. Conformational relation of Raman lines at 510 and 525 cm^{-1} with regard to the disulfide bond made by Sugeta et al. (1972, 1973) from studies of model compounds was confirmed in actual protein systems; 510 cm^{-1} to the gauch-gauch-gauch form of the disulfide bonds in lysozyme (Miyazawa and Sugeta 1974) and ribonuclease A (Miyazawa and Sugeta 1974) and 525 cm^{-1} to the trans-gauch-gauch form of the disulfide bond in lysozyme (Miyazawa and Sugeta 1974) and carboxypeptidase A (Chapter II of this dissertation). The 540 cm^{-1} S-S line was observed by Nakanishi et al. (1974) in their study of α-lactalbumin. The author assigns the Raman line at 540 cm^{-1} barely observed in the L. semifasciata toxin b spectrum to the trans-gauch-trans form of disulfide bond. This implies

that a certain number of disulfide groups in Laticauda toxin are in different geometrical conformations.

It was shown by Sugeta et al. (1972, 1973) that the C-S stretching frequencies depend upon molecular conformation about the C-C bonds adjacent to the C-S bonds. When there is a hydrogen atom at the trans site with respect to the C-S bond about the C-C bond, the stretching vibration lies at 630-670 cm^{-1} (P_H). Thus, the stretching vibration at 658 cm^{-1} may indicate that the conformation about the C-C bond is mostly in the P_H form. The line at 701 cm^{-1} in native P. platurus toxin which is weak but well defined may be due to either the C-S stretching of the methionine group or that of the P_N form. The spectrum of L. semifasciata toxin b powder in the C-S region (Figure 9) resembles that of P. platurus major toxin at 658 cm^{-1}, but does not show any peak near 700 cm^{-1} region. Since L. semifasciata toxin b does not contain any methionine residue, it is quite possible that the peak at 701 cm^{-1} in the spectrum of P. platurus major toxin is not due to the P_N form but due to the methionine residue. Thus, it may be assigned to the C-S stretching of the methionine group whose internal rotation about the CH_2-CH_2 and CH_2-S bonds in -CH_2-CH_2-S-CH_3 is in the trans-gauch form.

In Figure 8(b), a spectrum of heat treated P. platurus major toxin in solid is shown. The major change is the reverse of the relative intensities of the 850 and 832 cm^{-1}

lines. This spectral change indicates that the tyrosine residue has been more exposed. There is no change in the spectrum at the frequency attributed to the disulfide bonds (512 cm^{-1}). However, the line broadened and a definite shoulder appeared at 546 cm^{-1}, indicating that disulfide bonds were not broken by heat treatment, but they may have been slightly strained. The amide I band remained essentially unchanged. The amide III peak remained at 1243 cm^{-1} but broadened somewhat. It suggests that the heat denaturation of the toxin does not make any serious change in the peptide backbone conformation. This is reasonable since the four disulfide bonds hold the molecule rather tightly in a folded conformation. However, the slight change observed in the S-S Raman band and in the C-C-N, C-C, C-N bands suggest that some loops of the toxin molecule have been modified somewhat by heat treatment. The integrity of the four disulfide bonds maintained the structure of the peptide backbone so that the molecule could not unfold. The heat-treated toxin retained its original backbone structure with a slight twist. This relationship is diagramatically illustrated in Figure 10. This is consistent with the fact that the neurotoxin is relatively stable to heat treatment. It was found from studies of Tu et al. (1971) that L. semifasciata toxins retained toxicities even when boiled at 100°C for 30 minutes.

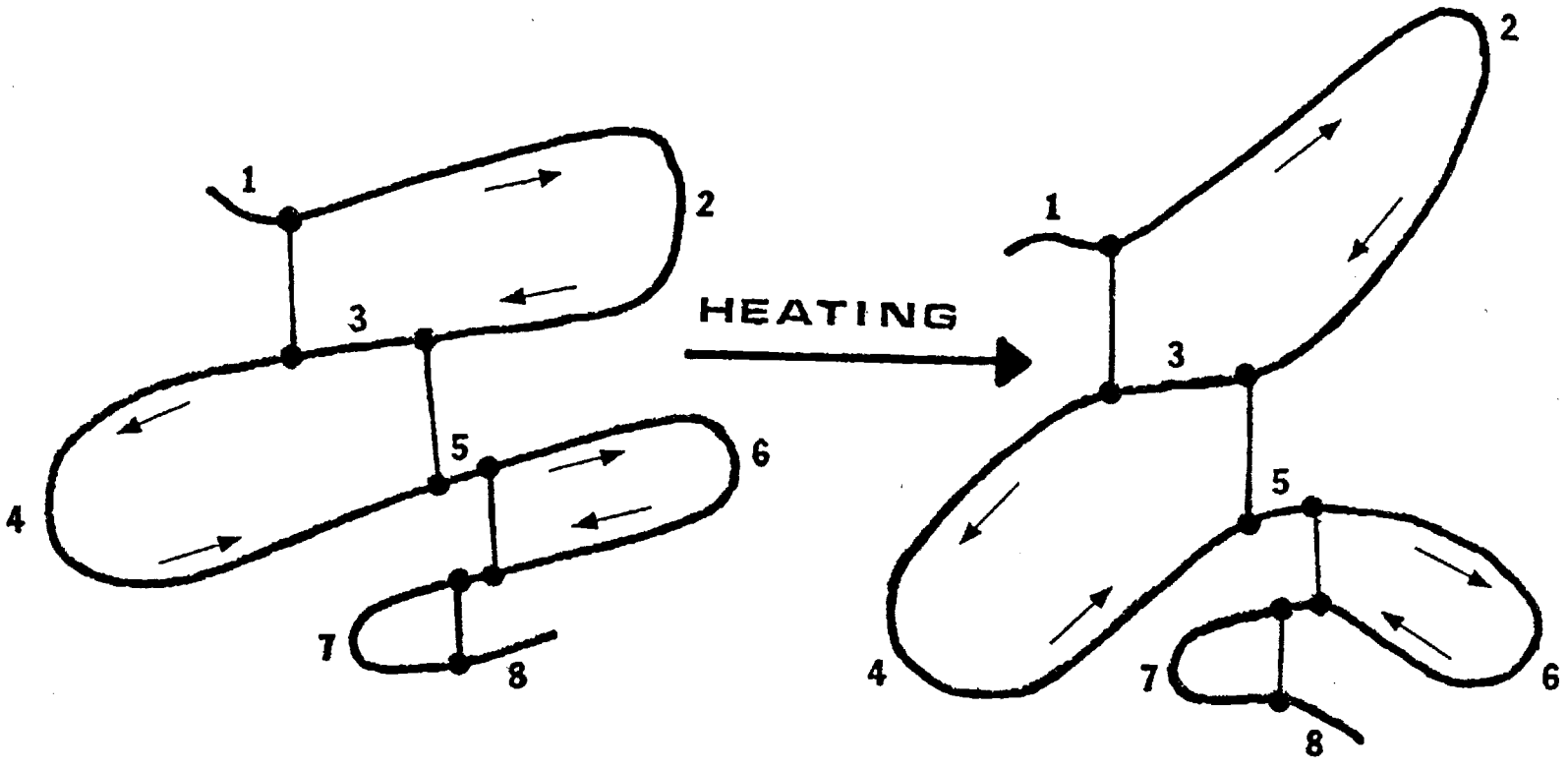

Figure 10. Proposed Backbone Conversion of Neurotoxin by Heat-treatment

CHAPTER IV

CONFORMATION OF INSULIN IN VARIOUS STATES AND BOVINE C-PEPTIDE AND THE PREDICTION OF THE STRUCTURES OF BOVINE AND PORCINE C-PEPTIDES

Introduction

In previous work on Raman spectra of native insulin in the solid state (Yu and Liu 1972, Yu et al. 1972, Yu and Jo 1973), samples in the crystalline powder form had been used for investigation. A typical method of handling such a sample for Raman scattering is to pack it into a conical cavity at the end of a metal rod. The laser beam is then focused and directed onto the sample at the grazing angle so that the scattering column is a strip on the powder surface (Yu et al. 1972). However, it is felt that the Raman spectra obtained under such conditions may not always be the true reflection of the structures of proteins in the crystalline states, depending on the laser power level at the sample, the relative humidity, and how the beam is focused. The problems involved are related to the loss of the order of protein crystals on air-drying and the thermal effect of the laser beam on the powder surface. A recent estimate from the measurements of the intensity ratios of the Stokes and anti-Stokes lines indicated that the

temperature of an air-dried powder sample under the heating of a well-focused beam of about 120 mW can be as high as 70°C or even higher. Even though the sample is placed in an environment with high humidity, the laser thermal effect may still cause the dehydration of the molecules and thus change the protein conformation. To avoid the uncertainites or complication, it was thought that the Raman spectra of native insulin should be reexamined by using large single crystals.

Recently, several perfectly shaped single crystals became available from Professor D. C. Hodgkin's group at Oxford. At the same time, Raman techniques in our laboratory have been improved so that I have been able to handle single crystals soaked in the mother liquor (Yu and Jo 1973b). The advantages of single crystals over powder for Raman scattering are obvious. First, a large, trnsparent single crystal permits the beam to travel through without Tyndall scattering and being appreciably absorbed. Secondly, since the single crystal is completely immersed in its mother liquor, there will be no dehydration effect. Furthermore, the concentration of protein molecules per unit volume is presumably greater in a single crystal than in the powder. Therefore, in a single-crystal Raman experiment one would expect an improvement in the quality of the data and that the spectra should truly manifest the protein structures in the crystalline state as determined by the X-ray diffraction

techniques.

Here, the author reports the single-crystal Raman spectra of insulin and deuterated insulin and present the true amide III contour. Spectral differences between the single-crystal and powder spectra will be interpreted.

The denaturation of globular proteins, insulin and lysozyme, has been studied by the Raman effect (Yu et al. 1972, Chen et al. 1973, 1974). In the case of insulin, only thermal denaturation was made. It is of interest to examine the effect of other denaturing agents on this hormone protein, both to determine their structural effects as compared with thermal denaturation and also try to learn something about the changes in the spectra of various important constituents of the molecule when their intramolecular environments alter. Since chemical denaturation can be carried out with a variety of agents, some measure of control over the conformational changes can be exerted and the spectroscopic changes produced thereby can be interpreted more precisely. In this paper, the effects on the Raman spectrum of insulin by trichloroacetic acid(TCA) and by S-S bond reduction and S-carboxymethylated are reported. Also, much improved spectra of fibril insulin (a thermally denatured product) in the "wet" and "air-dried" states are presented.

Raman spectra of C-peptide from bovine proinsulin is presented in Figures 10 and 11. The function of the

C-peptide in proinsulin is not clearly established, but it has been assumed that the C-peptide link is necessary in the correct formation of proinsulin prior to formation of the three disulfide bridges (Steiner and Clark 1968), it makes proinsulin biologically inactive by covering the biological active surface of insulin (Markussen and Schiff 1973) and it inhibits the aggregation by Zn^{++} ions of proinsulin hexamers to insoluble aggregates to facilitate the proteolytic conversion of proinsulin to insulin (Grant et al. 1972) in the membrane-limited granules. The C-peptide is produced together with two basic dipeptides or four basic amino acids when the connecting peptide is released in the conversion of proinsulin to insulin (Kemmler et al. 1971). The C-peptides show a great species variation in the amino acid constituent and chain lengths. The C-peptide from the bovine has 26 amino acids (Clark et al. 1969, Salokanggas et al. 1971, Steiner et al. 1971), porcine, 29 (Chance et al. 1968) and human, 31 (Oyer et al. 1970, Ko et al. 1971).

The circular dichroism(CD) spectrum of bovine C-peptide indicated that the gross structure is random-coil, but with a short segment of α-helix consisting of about six amino acids out of a total 26 (Markussen 1971). Frank and Veros (1968) showed the optical rotatory dispersion(ORD) and CD data of porcine proinsulin and insulin indicating some differences in α-helical content between two molecules. But they could not explain the difference only claiming an

assumption that the connecting peptide chain exists in a random-coil conformation. X-ray crystallographic studies (Fullerton et al. 1970) indicated, however, the presence of short segments of α-helices in the connecting peptide in crystals of bovine proinsulin. Recently, Markussen and Schiff (1973) has established an assumption from their hydrodynamic data that the C-peptide is in either a double stranded peptide (antiparallel) with the terminals positioned close to each other or a planar ring with the terminals closing the loop. In this paper, the author reports the Raman data of bovine C-peptide and the data obtained thereby is interpreted in terms of backbone conformation. To complement the experimental results, the author has computed the prediction of the bovine and porcine C-peptide backbone conformation using the method of Chou and Fasman (1974).

Materials and Methods

Single crystals of porcine zinc-insulin were grown by Mrs. S. M. Cutfield of Oxford according to the methods described by Schlichtkrull (1956). They were kept in the mother liquor comprising 0.006 M $ZnSO_4$, 0.05 M sodium citrate, 0.01 M HCl and 15% (V/V) acetone at pH 6.0-6.2. To deuterate the crystals, they were soaked in D_2O at room temperature for 2 weeks with twice changes of D_2O. Deuterium oxide, 99.7%, was obtained from Merk Sharp & Dohme of Canada

Limited. The air-dried powder of porcine insulin was also provided by Mrs. Cutfield.

TCA-treated insulin powder was prepared from Lilly's Ultralente U-80 insulin zinc suspension in the same way as that described in Chapter I of this dissertation. The S-carboxymethylated insulin powders were supplied by Schwartz/Mann Research Laboratories and used without further purification.

Samples of insulin fibrils were prepared by the method of Waugh (1944). 100 mg of insulin in the crystalline powder state was dissolved in 1 ml of H_2O and pH was adjusted with 0.04 N HCl solution to 2.42 and heated at 100°C for 45 minutes. The prepared standard gel was used for the Raman studies of the wet fibrils. One portion of this gel was air-dried for dried fibrils. For a complementary study of backbone conformation of β structure, the bovine lens was employed for Raman experiments. The eyes of calves were obtained from local abattoirs. Immediately after slaughter the eyes were removed and cooled. The lenses were taken out, washed with cold distilled water, and used for the Raman experiment within 5 hours. Since the capsule was not removed and the scattered light came directly from the interior part, the author believes that Raman spectra obtained represent the native, intact state of the lens.

Sample of bovine C-peptide was supplied by Dr. J. Markussen of NOVO Research Institute, Copenhagen, Denmark,

in the form of Na_4-C-peptide. Bovine C-peptide-d was prepared by dissolving the powder of C-peptide into deuterium oxide (10 mg/ml) and then freeze-drying it.

In a typical experiment with single crystals, a large (2 mm), extremely sharp, colorless, flat rhombohedral crystal of insulin was selected and transferred to a 5 mm-diameter glass vial filled with its mother liquid. The crystal in the vial was carefully oriented so that the incident beam entered a crystal face without deflection, i.e., the beam was perpendicular to the crystal face. The direction of the beam was normally deflected from the scattering column inside the crystal. The scattered light was collected by an f/1.1 lens on the entrance slit of a Spex 1401 double monochromator. All the spectra presented here were obtained with 514.5 nm line of an argon-ion laser (Coherent Radiation 52G) and at 4 cm^{-1} resolution. The Raman system and the optical arrangements have been described in detail in Chapter I and II of this dissertation.

The newly formulated protein predictive model of Chou and Fasman (1974) was applied to the conformational prediction of bovine and porcine C-peptides. The assignments of helical (α) and β-sheet potential (β) are given under each residue: H_α (strong helix former), h_α (helix former), I_α (weak helix former), i_α (helix indifferent), b_α (helix breaker), and B_α (strong helix breaker). Replacing the subscript α by β gives the respective β-sheet

potential assignments H_β, h_β, I_β, i_β, b_β and B_β. The rule of the α and β-forming nucleation followed here is that of Chou and Fasman (1974): when four helix formers out of six residues or three β-formers out of five residues are found clustered together in any native protein segment, the nucleation of these secondary structures begins and propagates in both directions until terminated by tetrapeptide breakers with 50% or more helix (or β-sheet) breaking or indifferent residues.

The relative probability that a tetrapeptide will form a β-turn is $p_t = (f_i)\ (f_{i+1})\ (f_{i+2})\ (f_{i+3})$ where f_i, f_{i+1}, f_{i+2} and f_{i+3} are respectively the frequency of occurrence for a certain residue at the 1st, 2nd, 3rd and 4th position of a β-turn. In a survey of 12 proteins, Chou and Fasman (1974) found that the average probability for any tetrapeptide to be in the β-turn is $p_t = 0.24 \times 10^{-4}$ and that $p_t = 0.5 \times 10^{-4}$ was chosen as a reasonalbe cut-off value in predicting β-turns. Using the f_i, f_{i+1}, f_{i+2} and f_{i+3} values from Table VII of Chou and Fasman (1974), the p_t values were computed for the 23 and 26 possible tetrapeptide combinations of the bovine and porcine C-peptides respectively and plotted in Figure 13.

Results and Discussion

Raman Spectra of Native Single-Crystal Insulin, Heat and Chemically Denatured Insulin and Reduced and Carboxymethylated A and B Chains of Insulin

Effect of Dehydration on the Raman Sepctra of Insulin.

In Figures 1 and 2, the author shows the Raman spectra of insulin at three specific orientations of a single crystal. The spectral features were found not to depend on the crystal orientation significantly, in agreement with the fact that the same functional groups in a protein crystal are more or less randomly orientad. The Raman spectrum of air-dried insulin powder obtained with the same spectral slit width (4 cm^{-1}) is included in Figure 1 for comparison. One readily sees that the apparent amide III contours of the top and bottom spectra in Figure 1 are somewhat different. The lines at 1245 and 1269 cm^{-1} have broadened and changed the relative peak intensities without drastically altering the peak positions. The characteristics of the amide III region of the Raman band of single crystal insulin may be explained on the basis of structural information obtained by X-ray studies (Blundall et al. 1971, 1972a, b). In the rhombohedral crystal, the unit cell contains six molecules of insulin. The crystal symmetry requires that this hexamer be organized as three equivalent dimers related to each other by three-fold rotation. The two-fold relationship between the molecules of dimer means that the extended C-terminal residues of the β-chin run antiparallel to each other. This

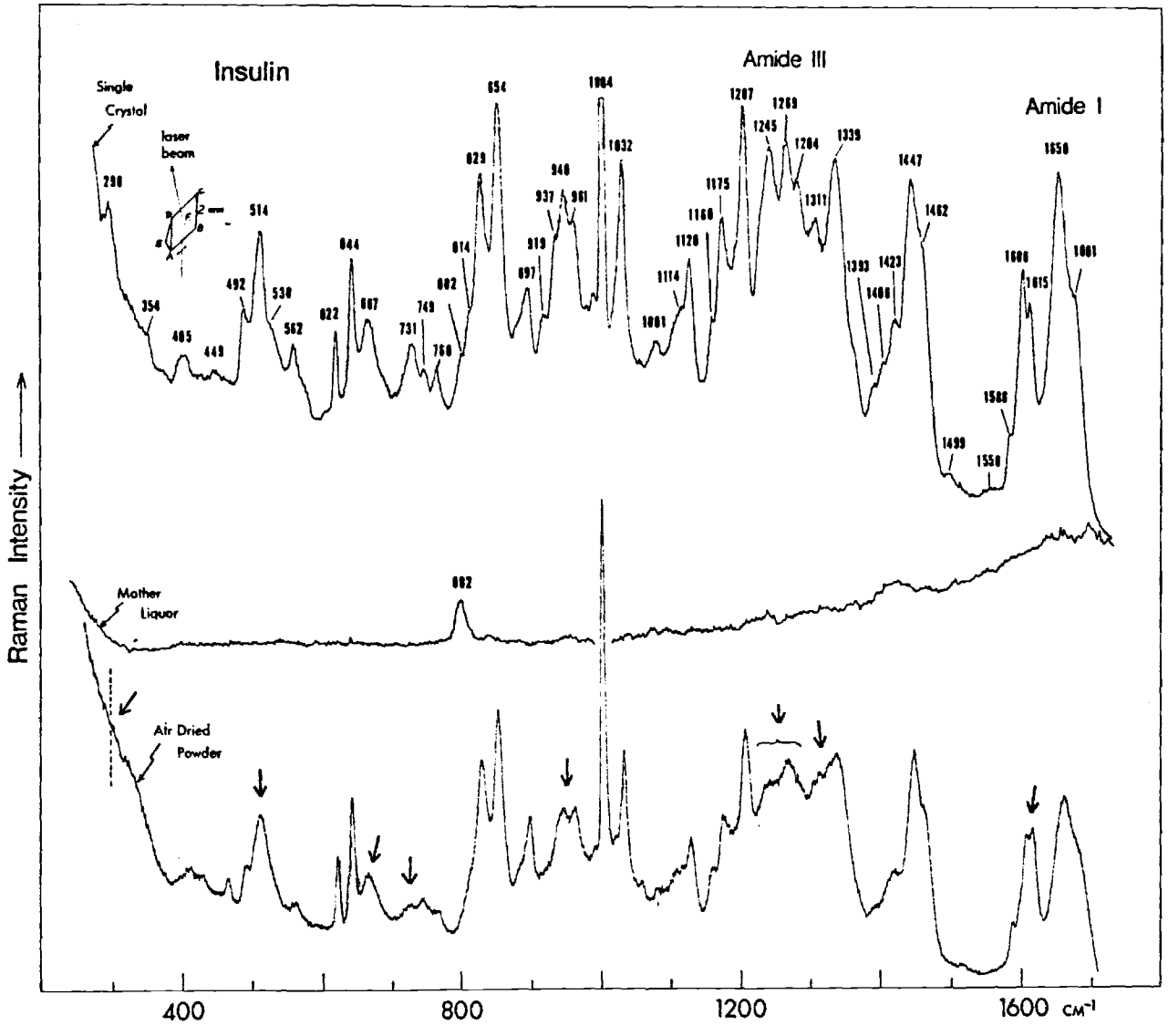

Figure 1. Raman Spectra of Single-Crystal Insulin (porcine) and Its Air-dried Powder

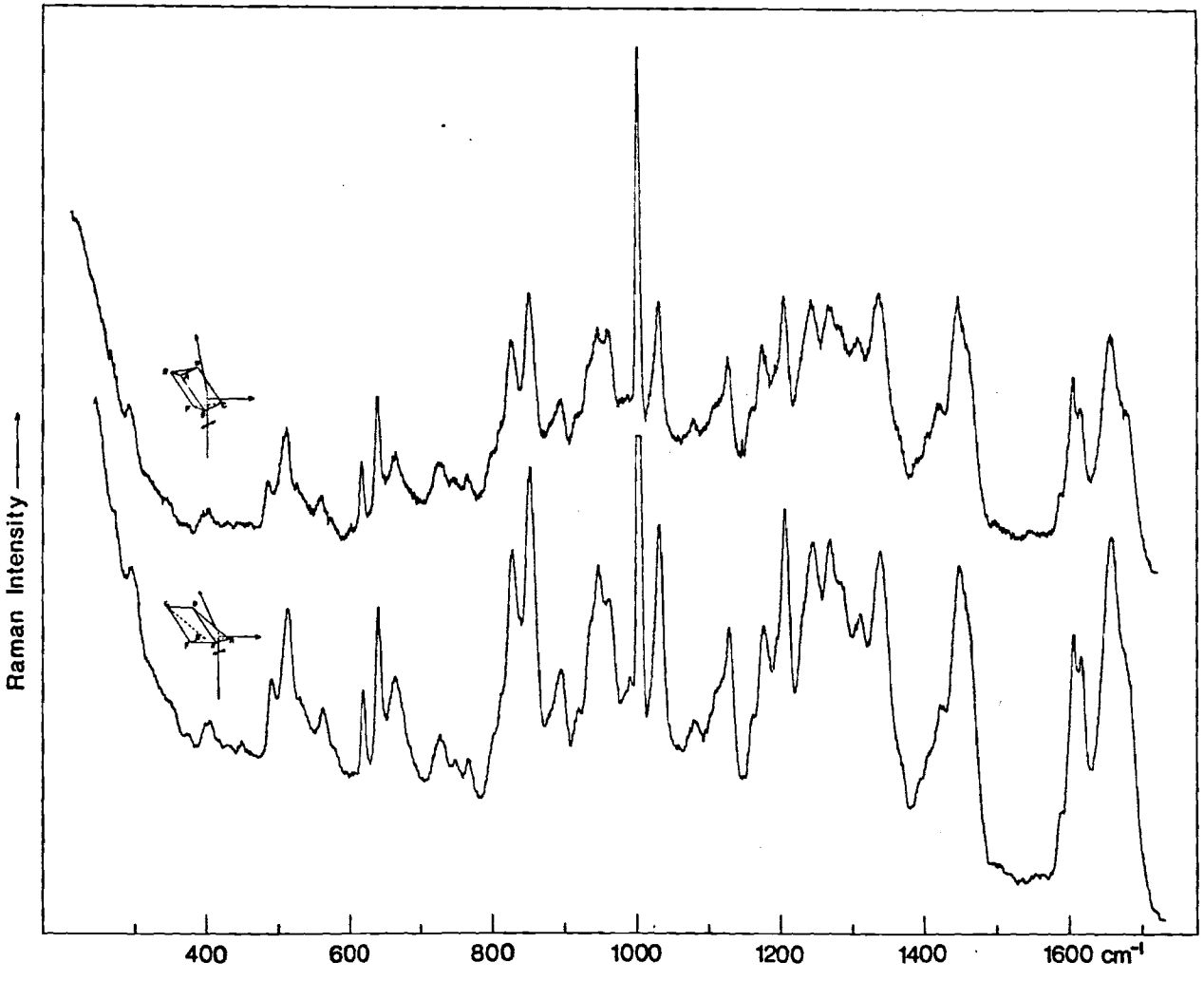

Figure 2. Raman Spectra of Insulin Single-Crystal (porcine) at Two Other Orientations

makes possible the antiparallel β-pleated sheet structure containing four hydrogen bonds between the monomers. These intermolecular hydrogen bonds connect the B24 and B26 phenylalanine peptide units of each monomer (Blundell et al. 1971, 1972a, b). The strength of hydrogen bonds of the antiparallel-β sheet and random-coil backbone are supposed to be more uniform in single-crystal state as reflected by sharpening of the line at 1245 cm^{-1} in the spectrum of single crystals. The side-chain interactions between monomers in the dimer and between dimers in the hexamer in the single crystals are believed to be different from that of the crystalline powder. One of indications of such interaction is revealed in the weak line at 449 cm^{-1} in the single crystal spectrum. Concerning this line, it was mentioned in previous paper (Yu and Jo 1973a) and Chapter I of this dissertation. Such kinds of different intermolecular interactions may also play a role in more ordering backbone of the single crystal.

Other spectral differences were also observed and are indicated by arrows in Figure 1. Of particular interest are the lines at 299 and 731 cm^{-1}. The 299 cm^{-1} line showed up in all the crystal spectra (Figures 1 and 2) and was not affected by deuteration (Figure 3). However, it disappeared (presumably too weak to see) from the powder spectrum. On the basis of the normal coordinate analysis on the C-S-S-C group (Yu 1969), one of the CSS bending vibrations was

predicted near 280 cm^{-1} if the dihedral angle of the C-S-S-C group is not far from 90^{o}. The author may tentatively assign the 299 cm^{-1} line in the crystal spectra to this type of vibration. The disappearance of this line on drying may be attributed to the change in the local geometry of the disulfide links. This behavior is quite similar to the marked intensity decrease of the C-S stretching at 667 cm^{-1} (see particularly Figure 4 of Yu et al. 1972) when the insulin molecules were incorporated into the cross-β-type fibrils (Burke and Rougvie 1972) with changes in the disulfide bond geometry (Yu et al. 1972). Other evidence for such structural change appeared in the broadening of the S-S line at 514 cm^{-1} and in the slight intensity decrease of the C-S line at 667 cm^{-1} (Figure 1). The proper explanation of the S-S stretching bands at 514 and 530 cm^{-1} could not be made in earlier Raman studies of insulin (Yu et al. 1972). The X-ray data of the cystine groups in the dimer state shows that two cystine groups (A7-B7 and A20-B19) are in the gauch-gauch-gauch forms and one intra disulfide linkage (A6-A11) in the gauch-gauch-trans form. The lines at 514 cm^{-1} in the single-crystal spectrum is assigned to two disulfide groups in the gauch-gauch-gauch form and the 530 cm^{-1} line to the one in the gauch-gauch-trans form. In the powder spectrum, the lines of the S-S stretching vibration are centered to one line at 514 cm^{-1} and broadened diminishing the shoulder at 530 cm^{-1}. This variation indicates that the

disulfide bonds are subtly distorted. The lines at 667 and 731 cm^{-1} may be assigned to the C-S stretching of the half-cystine groups in the P_H and P_C forms. Insulin crystal has four half-cystine groups of P_C form and two of P_H form (Blundell et al. 1972a, b). In crystal, the arrangement of the interchain disulfides of A7-B7 and A20-B19 which connect themselves at each end of their central α-helix (α helical regions, A2 to A8, A13 to A20 and B9 to B19) buries the cystine A6-A11. Since A6 and A11 are in the P_C and P_H forms respectively, the lines at 667 and 731 cm^{-1} are partially contributed by these two C-S stretching vibration. The increase of these lines on crystallization may be attributed to such local environmental change of the A6-A11 disulfide bond in the well organized single crystals. It is presumably possible to assume that the Raman lines arising from the C-S stretching vibration are affected not only by the number of the corresponding groups but also by their local environment. Since insulin molecule does not contain any methionine group, the C-S stretching region in the Raman spectrum of insulin may be sensitive mainly to the half-cystine groups in the molecule.

Single-Crystal Raman Spectrum of Deuterated Insulin and the True Amide III Contour. To develop Raman spectroscopy for quantitative measurement of the percentages of various structural components in a protein backbone, it is necessary to determine the contribution of side groups'

vibrations in the amide III region. In Figure 3 the author presents the Raman spectrum of insulin-d single crystal. The line at 1260 cm^{-1} is not due to "hard-to-exchange amide hydrogens" (Blout et al. 1961) since it could not be removed after 2 weeks' soaking in D_2O at room temperature and had no sign of intensity decrease even after 4 weeks. It also appeared in the spectra of insulin D_2O solution which was heated at 40°C for 60 minutes. According to the spectra of amino acid tyrosine solutions at acidic pH and pD values (Burke and Rougvie 1972), a characteristic tyrosyl ring vibrational mode is expected in this region. Since the intensity and frequency of this line in tyrosine is not appreciably affected by deuteration (Burke and Rougvie 1972), the author may consider the contour in the 1150 to 1350 cm^{-1} region of the deuterated spectrum (Figure 3) as the zero line for the amide III vibrations. Thus, the true amide III contour of insulin was obtained by graphical subtraction (see curve c, Figure 4). In Figure 4 the curves denoting (a) and (b) were redrawn from the top spectrum of Figure 1 and that of Figure 3, which were both recorded by adjusting the 622 cm^{-1} (due to Phe) line intensity to the same value. Since the phenyl rings do not have acidic hydrogens, the intensity of this line should not be altered by deuteration. However, this may not be so with the tyrosyl lines because the phenolic hydrogen can be substituted by deuterium. The intensity of the 854 cm^{-1} line (relative to

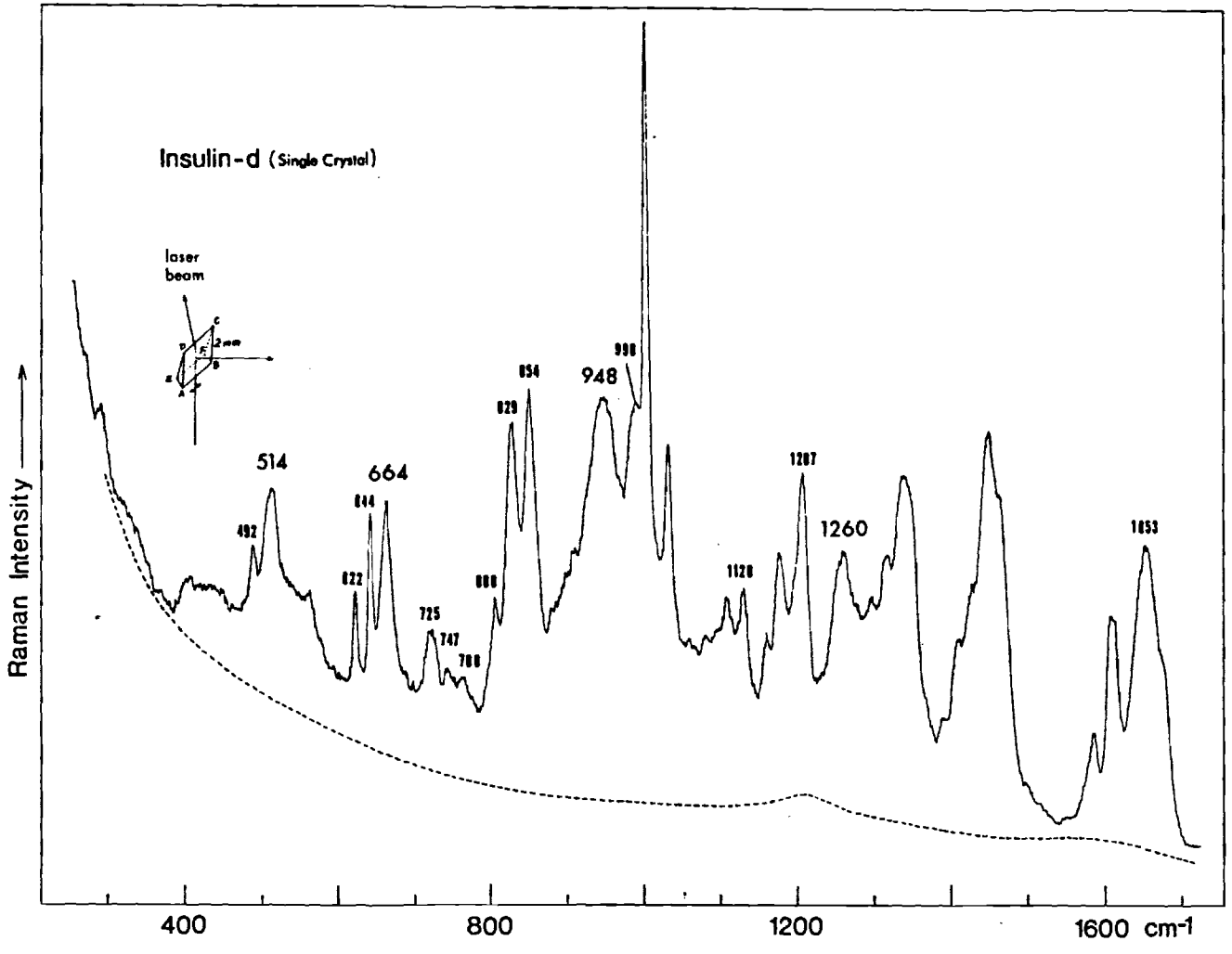

Figure 3. Raman Spectrum of Single-Crystal Insulin-d (porcine)

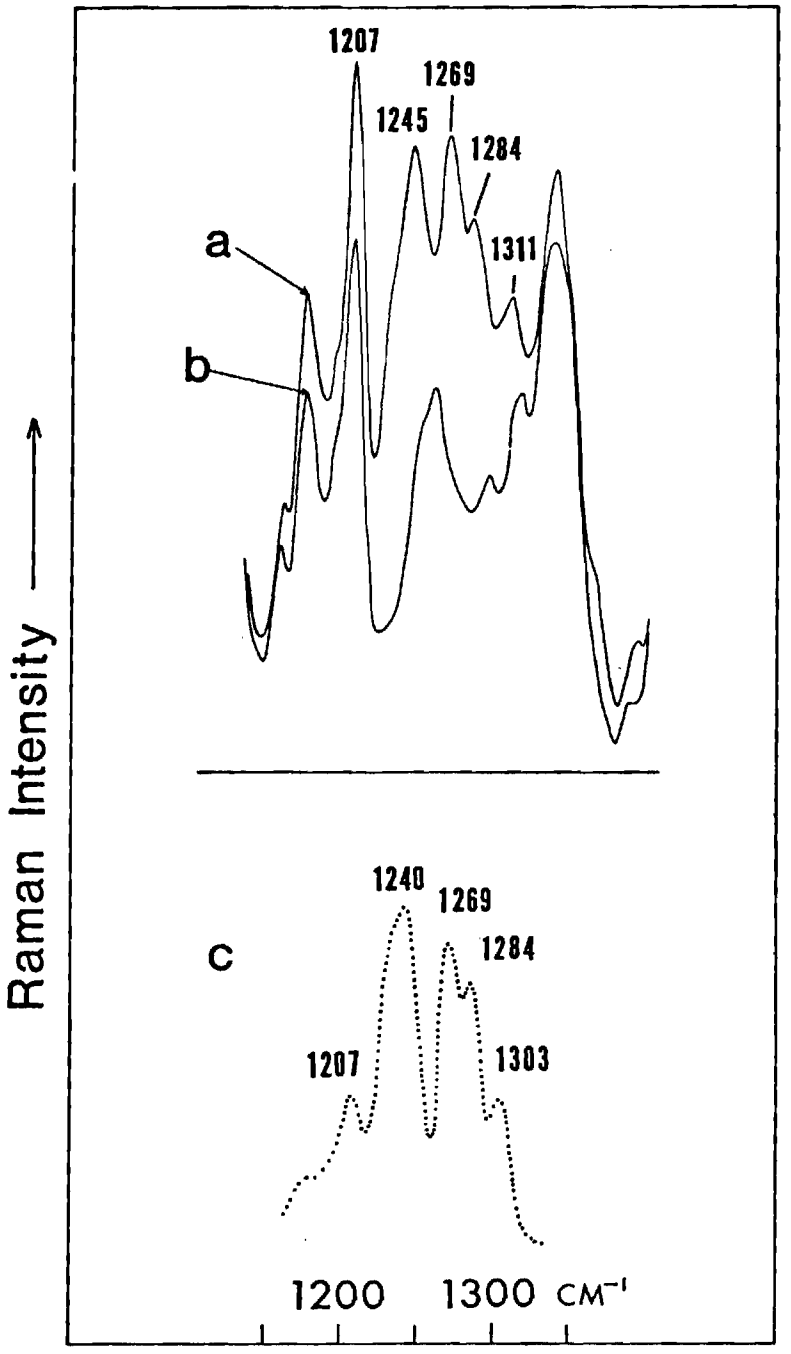

Figure 4. Difference between Spectra of (a) Insulin and (b) Insulin-d and (c) Its Difference Spectrum in Amide III

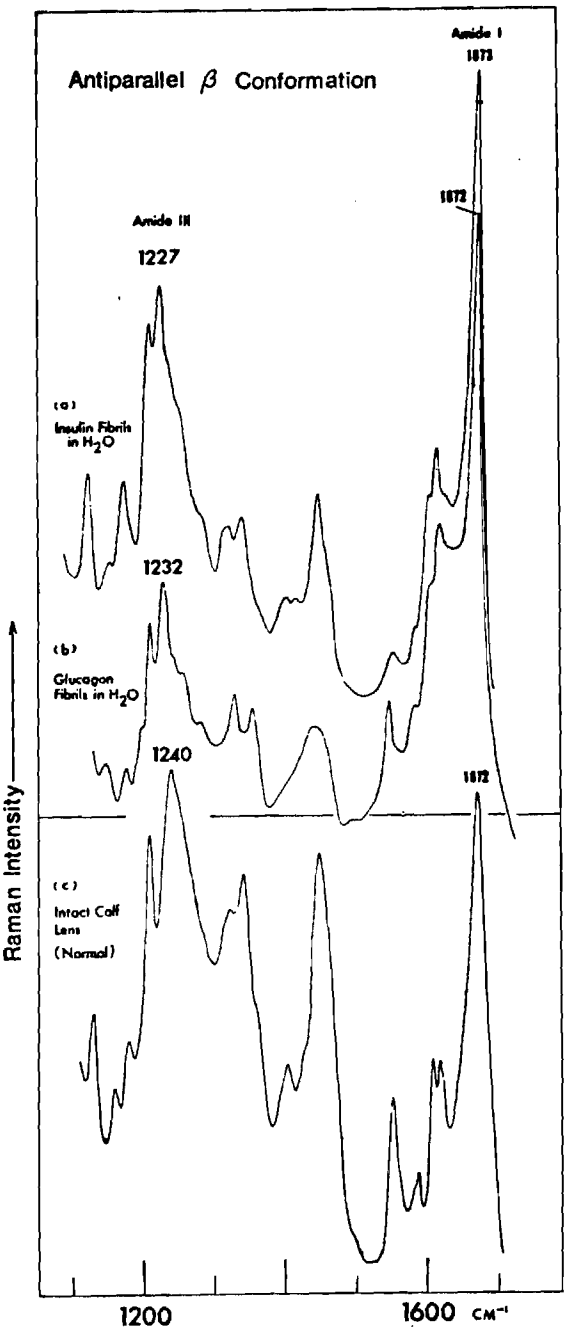

Figure 5. Raman Spectra of Three Proteins containing Antiparallel-β Conformation

the one at 829 cm^{-1}) is, therefore, somewhat reduced on deuteration. The tyrosyl line at 1207 cm^{-1} may also have been affected, resulting in a small peak at 1207 cm^{-1} in the difference spectrum (curve c of Figure 4). This line should be discarded as far as the true amide III contour is concerned.

Two aspects of Figure 4 deserve mention. First, the appearance of the true contour in (c) is quite different from that of the apparent one in (a). Secondly, the peak positions at 1245 and 1311 cm^{-1} in (a) have shifted to 1240 and 1303 cm^{-1}, respectively. This means that the frequencies measured in the untreated data may not always represent the true values of molecular vibrations.

The assignments of the various components corresponding to 1240, 1269 and 1284 cm^{-1} have been given by Yu *et al.* (1972). They assigned the line at 1269 and 1284 cm^{-1} to the α-helical form and the one at 1240 cm^{-1} to the random-coil (non-H-bonded or weakly H-bonded) with a partial contribution from the antiparallel-β structure in the formation of insulin dimer in crystals (Blundell *et al.* 1971). At present, the author wishes to maintain these assignments, but adding the newly discovered line at 1303 cm^{-1} to the α-helical category. The splittings of the α-helical amide III vibrations may represent different strength of H-bonds or due to the irregularity of α-helical structure.

The assingments of the amide III vibrations to various

structural features in protein sometimes cannot be made with certainty without information from the X-ray crystallographers. The reason Yu et al. (1972) assigned the line at 1240 cm^{-1} to the random-coil form predominantly is because a fully solvated (H-bonded) random-coiled polypeptide (e.g., poly-L-glutamic acid at pH 10, and freshly prepared glucagon solution at pH 2) (Lord and Yu 1970a, Yu and Liu 1972) has its Raman amide III vibration at 1238 cm^{-1}. Since the vibrational interactions between peptide groups in the unordered component are on the average zero (Krimm 1962) the strength of H-bonding may become important in determining the amide III frequency for the random-coil structure. In principle, a weaker hydrogen-bond gives rise to a lower bending frequency (amide III) (Richards and Thompson 1947). It is quite conceivable that those peptide units not involved in α helices and β structure may be only weakly H-bonded or non-H-bonded at all. Therefore, the assignment of the 1240 cm^{-1} line to such structure seems reasonable. On the other hand, as the author will discuss in the section (d), a Raman line at 1240 cm^{-1} may also be assigned to the antiparallel-β structure. Since it is already known from the X-ray crystal structure of insulin (Blundell et al. 1971, 1972a, b) that there are only four peptide units involved in the antiparallel-β form for dimer formation, the assignment of the 1240 cm^{-1} to β structure exclusively can be eliminated. As to the assignments of α helices, Yu et al. (1972) has used

the α-helical poly-L-alanine (Fanconi et al. 1969) and crystalline glucagon as model compounds (Yu and Liu 1972).

Effect of Deuterium Substitution on the Geometry of the C-S-S-C Group. Upon deuteration, there appears an unusual intensity increase of the C-S line at 667 cm^{-1} (Figure 3). In order to find out the nature of this intensity alteration, it may be necessary to consider various possible causes. Since the C-S-S-C group is separated from the nearest peptide units by the α carbons and the force constant of the C-S bond is quite different from those of the C-C and C-N bonds, the C-S-S-C group vibrations are not expected to couple to an appreciable extent with neighboring peptide units. Deuteration of the peptide groups should not result in drastic changes in either frequencies and intensities of the S-S and C-S stretching vibrations. This expectation was actually confirmed in the case of lysozyme (Lord and Yu 1970a). Another possibility is the effect of D_2O substitution for H_2O solvation of the exposed disulfide bond (B7-A7) (Blundell et al. 1971), because it may affect the electronic structure of the chemical bonds concerned and thus altering the differential polarizability. To test this, the author checked the Raman spectra of cystine in H_2O and D_2O (Yu 1969) and found that the intensity ratio of the C-S to S-S lines is not changed by D_2O solvation. In view of the fact that in most crystals the D-bond is slightly larger (Pimental and McClellan 1960) and stronger (Lewin and

Rosmarin 1970, Lewin 1969) than the H-bond, it is quite conceivable that deuterium substitution in protein crystals may distort the local geometry (P_H forms in more uniform) or change the local environment of the disulfide links, resulting in the marked intensity increase at 667 cm^{-1}.

The Raman Amide I Mode and the Variations of the Amide III Frequencies in Antiparallel-β Conformation. The theoretical basis for understanding the amide I vibrations of various polypeptide-chain conformations was first provided by Miyazawa (1960). Based on his perturbation treatments, it was predicted that the strong infrared bands of both antiparallel- and parallel-β structures were at nearly identical frequencies and that the totally symmetric $\nu(0,0)$ mode for polyglycine I should appear in the Raman spectra at 1648 cm^{-1}, which is quite different from the actually observed value at 1674 cm^{-1} (Small et al. 1970). Recently, in an attempt to achieve a more consistent understanding of splitting in the amide I modes of polyglycine I, Krimm and Abe (1972) modified the Miyazawa theory by including the transition dipole coupling term and neglecting the intrachain interaction between peptide groups, which was overemphasized in the original treatment. The modified theory immediately leads to the fact that the equations for the two types of β structures are different. The important consequence of their analysis is that the frequencies of the parallel-β sheet overlap those of the

α helix (approximately at 1657 cm^{-1}) and are distinguishable from those of the antiparallel-β sheet. For the antiparallel-β-sheet structure, the following relations hold:

$$\nu(0, 0) = 1674 = \nu_o + D_{ol} + D_{ll}$$

$$\nu_{\parallel}(0, \pi) = 1685 = \nu_o - D_{ol} - D_{ll}$$

$$\nu_{\perp}(\pi, 0) = 1636 = \nu_o + D_{ol} - D_{ll}$$

$$\nu_{\perp}(\pi, \pi) = 1723 = \nu_o - D_{ol} + D_{ll}$$

The value of $\nu_{\perp}(\pi, \pi)$ was predicted on the basis of $\nu_o = 1679.5$, $D_{ol} = -24.5$, and $D_{ll} = 19$ cm^{-1}. For the definitions of these symbols, the reader is referred to the paper by Krimm and Abe (1972). All four frequencies are Raman active but only the $\nu(0, 0)$ mode at 1674 cm^{-1} is expected to be "Raman intense". In addition to polyglycine I, this strong Raman line has actually been observed in poly-L-lysine (Wallach et al. 1970), insulin carboxymethylated A- and B-chains (see Figures 8 and 9), insulin fibrils (Yu and Liu 1972, Yu et al. 1972), glucagon fibrils (Yu and Liu 1972), and intact calf lens (see Figure 5). It should be noted that the present conclusion about the antiparallel-β structures of insulin and glucagon fibrils disagree with the suggested models (Rougvie and Shriver 1973, Burke and Rougvie 1972). They interpreted the X-ray diffraction data in terms of parallel pleated sheets.

Raman spectra of "wet" insulin fibrils, "wet" glucagon fibrils and intact calf lens are shown in Figure 5. One interesting aspect of these spectra is that they all show a strong single Raman amide I at 1673 ± 1 cm^{-1}, but the amide III frequencies vary in the range 1227-1240 cm^{-1}. Based on the assumption that the theory of Krimm and Abe (1972) is applicable to the much less regular heteropolypeptide chains of proteins, the frequencies at 1673 ± 1 cm^{-1} do suggest that these biological materials may contain exclusive antiparallel-β structure. The large frequency variations in the amide III region indicate that the bending vibrations of the N-H bonds in polypeptide chains are sensitive to some other structural parameters. One possible factor may be related to the strength of H-bonds. On the basis of the principle (Richards and Thompson 1947) that a weaker hydrogen bond gives rise to a lower amide III frequency within a given type of conformation (α helix, antiparallel-β, parallel-β, or random-coil), the lowest amide III frequency at 1227 cm^{-1} in Figure 5 means that the H-bonds in insulin fibrils are the weakest among these three materials. Further studies on this point seem desirable.

Effect of Water on the Structure of Insulin Fibrils. Water is known to affect the conformation of proteins (Yu and Jo, 1973b). In the work of Yu et al. (1972), attempts were made to detect the spectral differences between the "wet" and "air-dired" insulin fibrils (a thermally

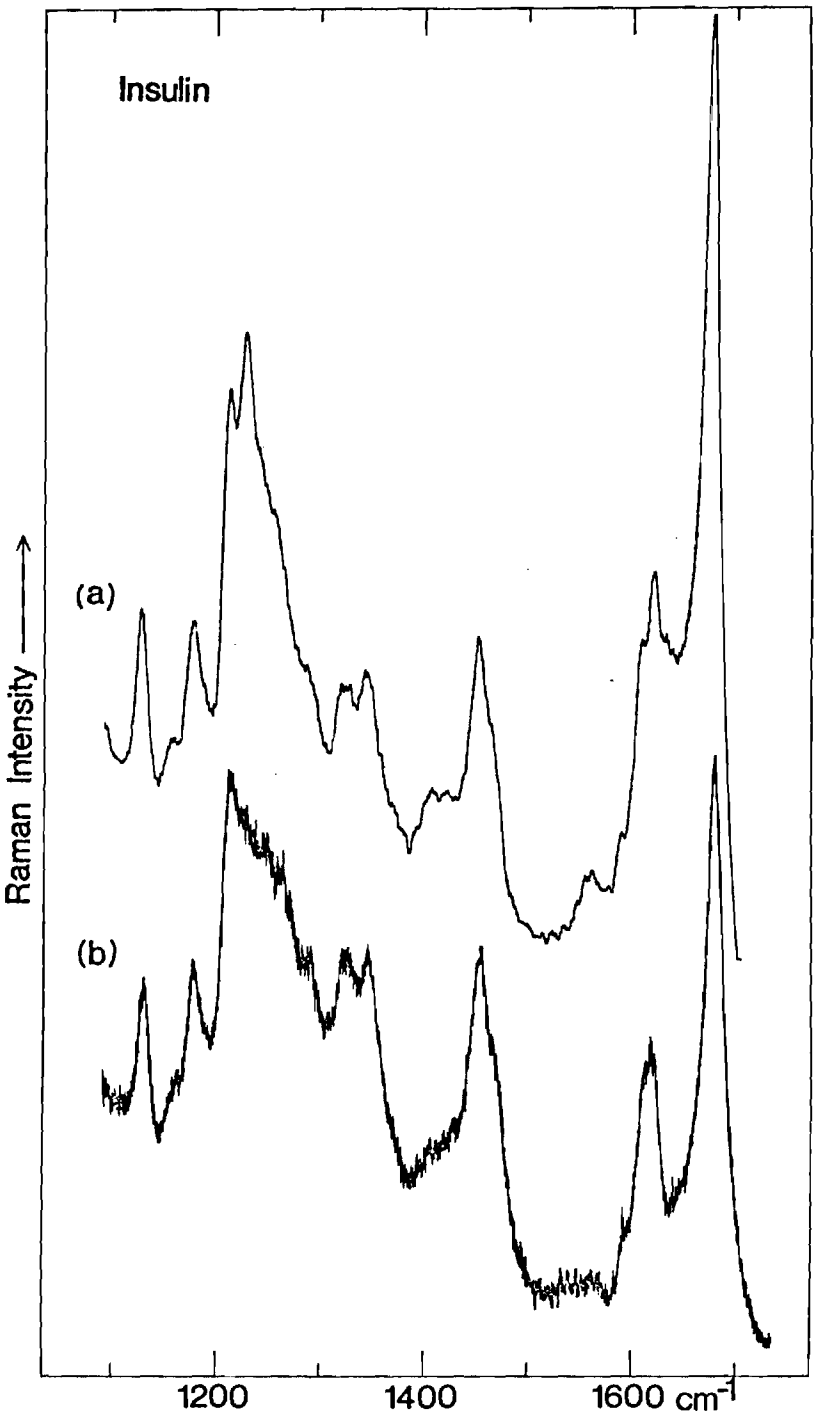

Figure 6. Raman Spectra of (a) "Wet" and (b) "Dried" Insulin Fibrils

denaturated product). Because of the experimental difficulties with the "wet" gel samples, they were not able to pinpoint the effect of dehydration although they did conclude that the β-structure of insulin fibrils remained in both "wet" and "dried" samples.

The recent improvement in Raman techniques has enabled the author to obtain a better result. In Figure 6 the author shows the Raman spectra of fibrillar insulin in the "wet" and "dired" states. The effect of dehydration is clearly seen in the amide III region, i.e., the sharp peak at 1227 cm^{-1} has broadened, indicating that the uniformity of H-bonds may be lost upon air-drying (Lord 1971).

Destabilization Effect of Trichloroacetic Acid(TCA). The Raman spectrum of TCA-treated insulin in the powder form (Figure 7) is compared to that of native state obtained by Yu et al. (1972). Changes in peak intensity and shape are noticed at several regions. The disulfide stretching at 515 cm^{-1} becomes sharper with the appearance of a shoulder at 530 cm^{-1} which has been also observed in the spectrum of heat denatured insulin (Yu et al. 1972). The C-S stretching line of half-cystine groups in the P_H form at 668 cm^{-1} has shifted to the lower frequency and decreased while a well-defined band at 741 cm^{-1} has appeared. This result may be an indication that a number of the half-cystine groups were transformed into a new form such as P_C. This assumption is possible since a dramatic change in amide III may sugggest

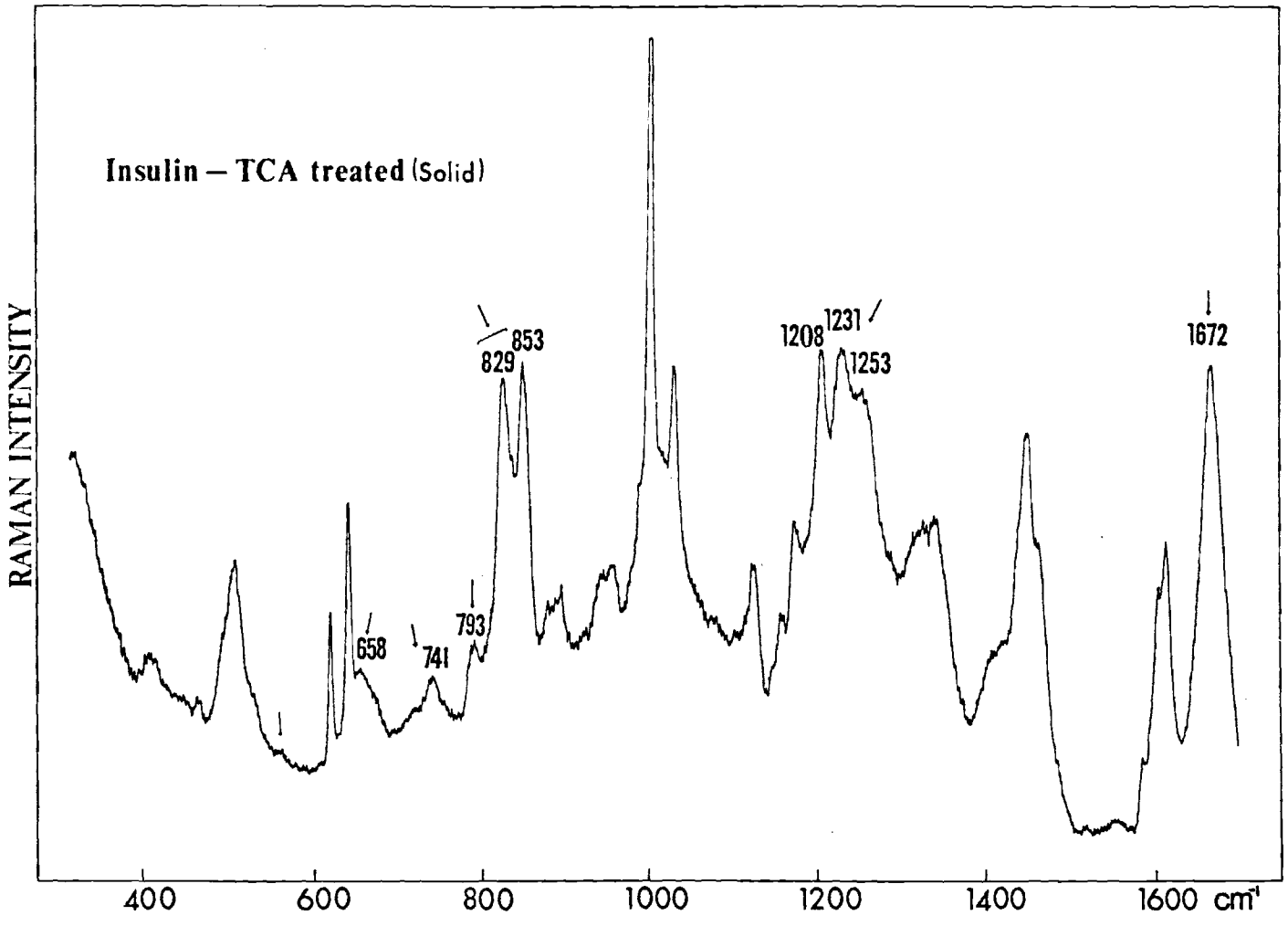

Figure 7. Raman Spectrum of TCA-treated Insulin Powder

an unfolding of backbone conformation. The characteristics of the S-S and C-S stretching lines can be understood when the amide I and amide III bands are examined. The disulfide links, A7-B7 and A20-B19, connects the groups of helical end-sides of each chain (Blundell et al. 1972a, b). On denaturation, a main-backbone conformation change is revealed in the amide III region. The 1270 cm^{-1} line of α helices disappeared or was masked and the two peaks 1231 and 1253 cm^{-1} are shown clearly. It may be a result of a transformation of α-helical segments in the backbone conformation into antiparallel-β and/or random-coil conformation. But as shown in Figure 4, the amide III contour contains an intense peak presumably arising from tyrosyl ring vibration at 1250 cm^{-1}. Therefore, the line at 1253 cm^{-1} may be eliminated as a backbone conformation indicator. The author assigns the 1231 cm^{-1} line to the antiparallel-β sheet. The amide I band at 1672 cm^{-1} supports this assignment. The intensity ratio of tyrosyl ring bands at 829 and 853 cm^{-1} in the spectrum of the TCA-denatured insulin is different from that of native one. Since the tyrosine residues are involved in monomer-monomer (B16 and B26) and dimer-dimer contacts (A14) in 3-fold native insulin crystals (Blundell et al. 1972a, b), TCA denaturation may be expected to affect these contacts. Other noticeable changes are in the regions of 563 and 880 to 1200 cm^{-1}. It is reasonable to anticipate the change in C-C, C-N and C-C-N stretching region of the

spectrum since it should be sensitive to the backbone conformation of the protein.

Raman Spectra of Reduced and Carboxymethylated A- and B-Chains of Insulin. In Figure 8, the Raman spectrum of carboxymethylated A-chain is presented. The A-chain does not contain any phenylalanyl residues, which is well demonstrated by not showing any signals at 622, 1006, 1032, 1588 and 1607 cm^{-1}. The Raman band at 702 cm^{-1} may be due to C-S stretching of half-cystine groups, while the 763 cm^{-1} band is due to the newly formed C-S bond between the sulfur atom of half-cystine group and the carbon atom of methyl group in $-CH_2COOH$. This line is also observed in the spectrum of carboxymethylated B chain. Presently it is impossible to derive any information about the geometry about the C-S bond. The amide III and amide I bands show a typical type of the antiparallel-β structures of the backbone conformation. A small peak at 1251 cm^{-1} may be due to one of the characteristic modes of tyrosine groups. The shapes of the amide III and amide I are rather close to those of insulin and glucagon fibrils shown in Figures 5 and 6. The spectrum of carboxymethylated B chain shown in Figure 9 has an amide I band somewhat broader than that of carboxymethylated A chain. This fact may indicate that the hydrogen bonds of the antiparallel-β structure present in the carboxymethylated B chain are less uniform.

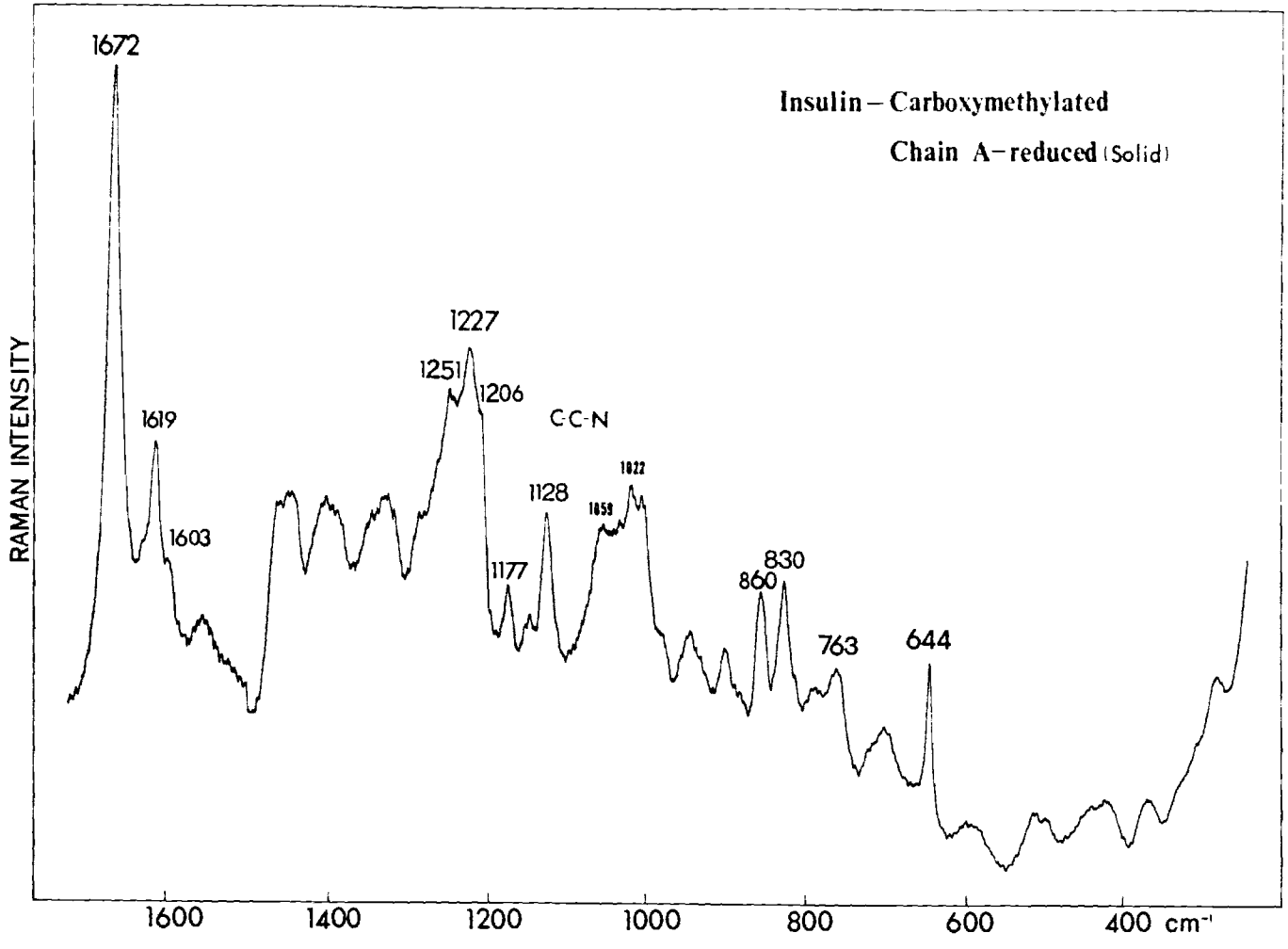

Figure 8. Raman Spectrum of Reduced and Carboxymethylated Insulin Chain A

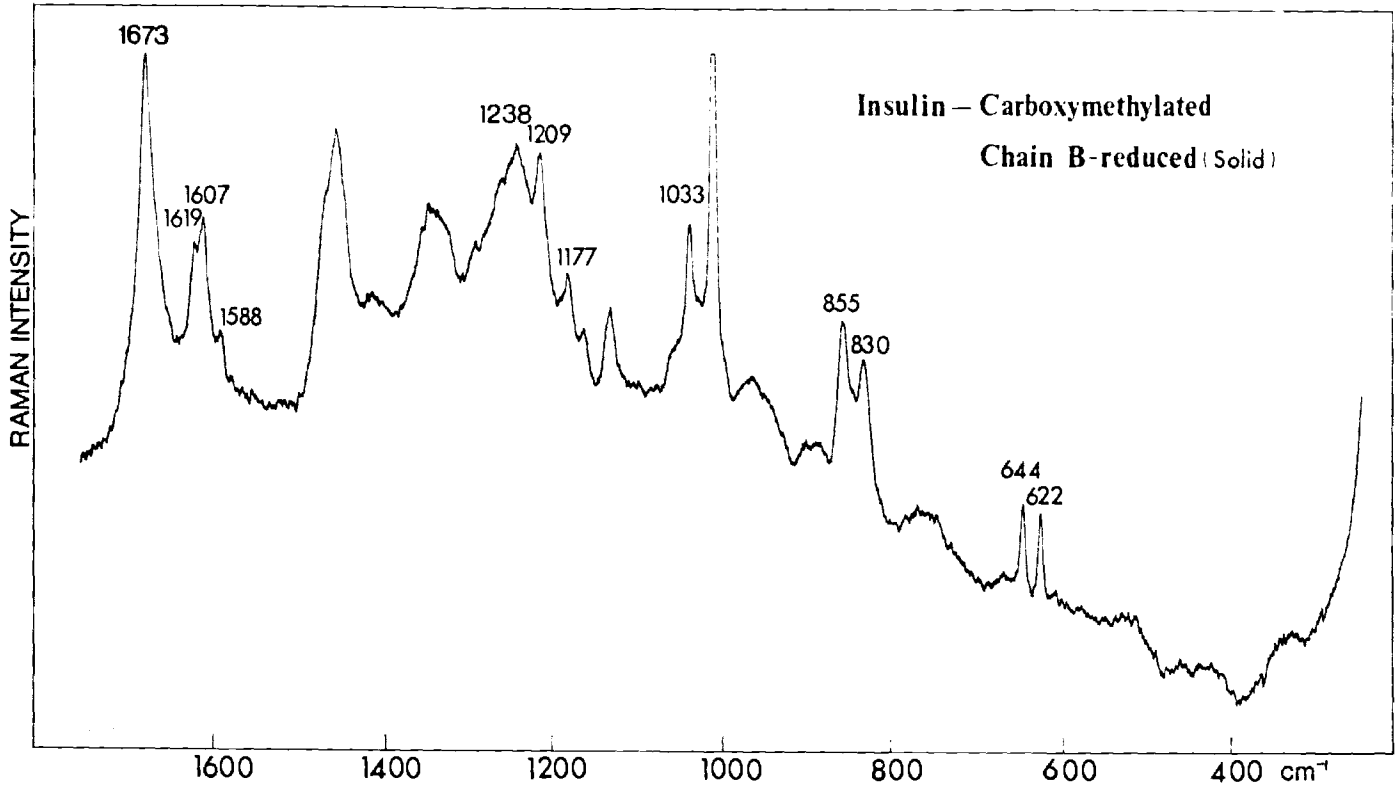

Figure 9. Raman Spectrum of Reduced and Carboxymethylated Insulin Chain B

Raman Spectra of Bovine C-peptide and the Prediction of the Structures of Bovine and Porcine C-peptides

The bovine C-peptide does not contain any aromatic side groups (Clark et al. 1969, Salokanggas et al. 1971, Steiner et al. 1971) and this fact is well reflected on the spectra in Figures 10 and 11 showing no Raman signals characteristic of Trp, Tyr, Phe and His. The amide III region consists of three peaks at 1248, 1272 and 1280 cm^{-1}. The line of 1272 cm^{-1} may be due to another vibrational mode other than the N-H in-plane bending vibration, since the peak remains after the sample is deuterated as shown in the spectrum of bovine C-peptide-d. The amide I band is distinctly dissimilar in frequency and shape to that of antiparallel-β structure. Therefore, the line at 1248 cm^{-1} may be assigned to a high portion of random-coil structure and 1280 cm^{-1} to α-helical segments. In the spectrum of the deuterated C-peptide, the amide I band has shifted to the lower frequency due to mass effect. Centrifugal and CD data of proinsulin and insulin were reported by Frank and Veros (1968). They concluded that the insulin moiety of proinsulin exists in the same conformation as in insulin itself and attributed the difference in CD data between the two molecules which indicated different contents of α helices to the probable random-coil structure of the connecting peptide. Their assumed random-coil structure of the C-peptide was based on predominant number of residues favoring a disordered

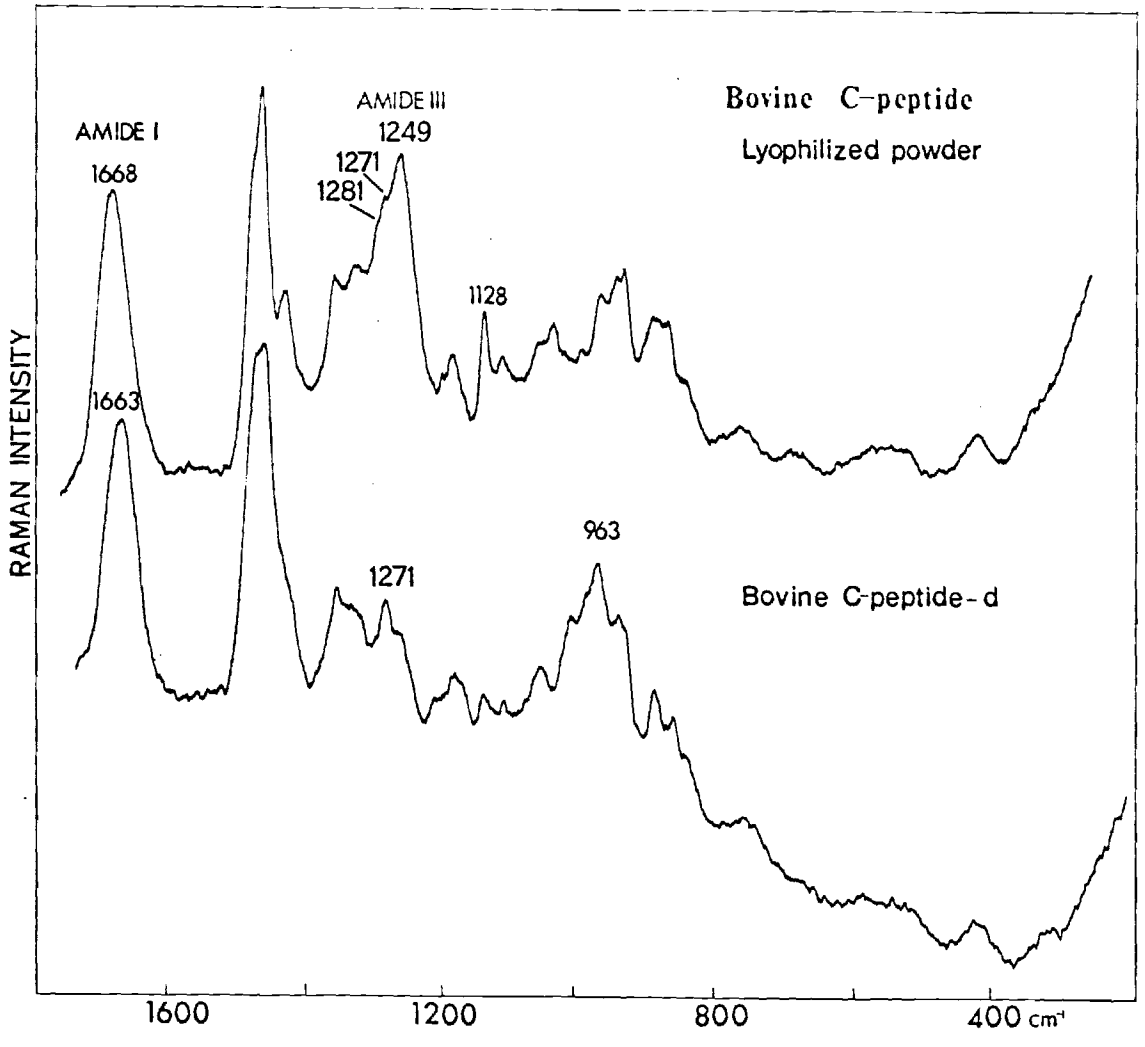

Figure 10. Raman Spectra of C-peptide (bovine) and C-peptide-d in Solid

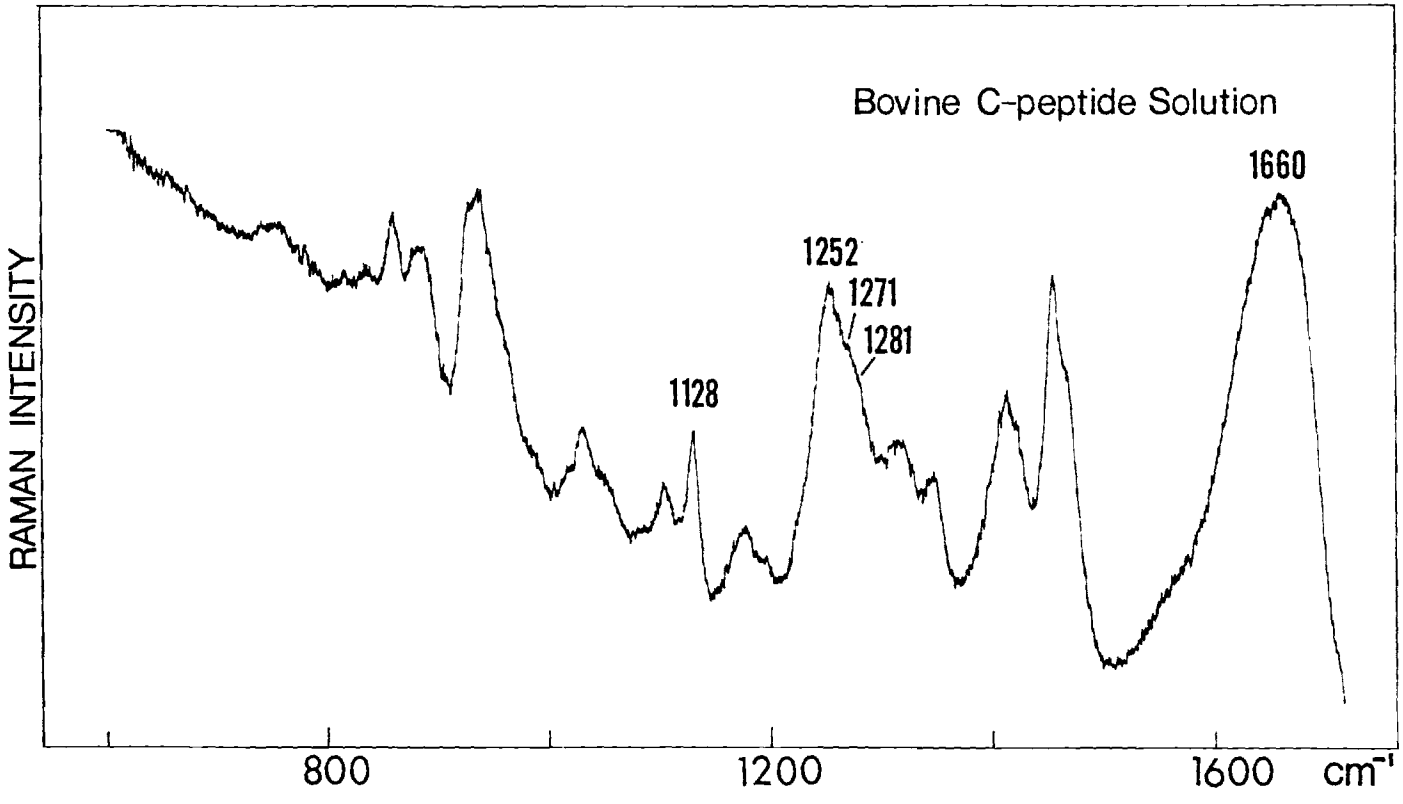

Figure 11. Raman Spectrum of C-peptide (bovine) in Solution

structure, proline and glycine. From the Raman result reported here, it seems clear that the difference in CD spectra between proinsulin and insulin was not due to the random-coil conformation of the C-peptide but due to the α-helical segments of the C-peptide. Actually the same conclusion was made by Yu et al. (1972) from their Raman studies of porcine insulin and proinsulin, in which they interpreted the spectrum differences as being a contribution from the considerable fraction of α-helical structure. However, X-ray crystallographic studies (Fullerton et al. 1970) indicated the presence of short segments of α helices in the connecting peptide in crystals of bovine proinsulin. Similar results were obtained by Markussen (1971) using CD spectrum and predicted that six amino acids may be located in that region. The Raman spectrum of C-peptide in aqueous solution is shown in Figure 11, which is similar in the amide I and III regions to that of lyophilized powder (Figure 10).

Recently, Chou and Fasman (1974) have developed a new technique in predicting protein conformation with about 80% overall accuracy. The author used this calculation technique to predict both bovine and porcine C-peptide conformation and compared with the experimental results. It has already been concluded that porcine C-peptide in proinsulin may have a considerable α-helical fractions by Yu et al. (1972) from their studies of the difference Raman spectrum between

proinsulin and insulin.

The conformational regions of the bovine and porcine C-peptides computed by the above method are summarized in Table 1. The regions of random-coil conformation (C) have both $<P_\alpha>$ and $<P_\beta>$ values less than unity except residues 1 to 5 of both bovine and porcine. If this portion is counted as an α region, following the condition A-1 stating that four helical residues out of six residues is necessary for helix nucleation, Gln 6 should be included, and then Pro 5 is located in the inner helix. This violates the condition A-3. Moreover, the facts that Pro cannot occur at the C-terminal helical end (condition A-3) and Pro prefers the N-terminal helical end (condition A-4) are against helix formation of residues 1 to 5. Helix termination at Gln 6 and Ala 13 in bovine C-peptide is made following the condition A-2 since the tetrapeptides, 2 to 5 (h_2b_2), and 14 to 17 (b_4) are enough to stop helix propagation having $<P_\alpha> < 1.00$. In Table 1, the bovine C-peptide is predicted as 31% helix and 69% coil. The CD studies of Markussen (1971) found the bovine C-peptide contains six residues (23%) in helix region. Together with X-ray finding of helical fractions of the connecting peptide in bovine proinsulin, these data support the finding of α-helical region in bovine C-peptide by Raman experiment. The porcine C-peptide is predicted to be 48% helix and 52% random-coil by the similar procedure used in the prediction for bovine

Table 1. Conformational Prediction for Bovine and Porcine C-peptides

Bovine C-peptide

Residues	Predicted[a]	$\langle P_\alpha \rangle$ [c]	$\langle P_\beta \rangle$ [c]
1 - 5	C	1.06	0.72
6 - 13	α	1.24	1.04
14 - 26	C	0.80	0.80

Predicted β-turns[b]	$p_t \times 10^4$ [b]	$\langle P_t \rangle$ [c]	$\langle P_\alpha \rangle$ [c]	$\langle P_\beta \rangle$ [c]
15 - 18	1.296	1.37	0.78	0.80
17 - 20	0.910	1.40	0.76	0.85

Porcine C-peptide

Residues	Predicted[a]	$\langle P_\alpha \rangle$ [c]	$\langle P_\beta \rangle$ [c]
1 - 5	C	1.17	0.55
6 - 12	α	1.23	1.02
13 - 18	C	0.67	0.88
19 - 25	α	1.37	1.01
26 - 29	C	0.72	0.82

Table 1. (Continued)

Predicted β-turns[b]	$p_t \times 10^4$[b]	$\langle P_t \rangle$[c]	$\langle P_\alpha \rangle$[c]	$\langle P_\beta \rangle$[c]
13 - 16	0.754	1.39	0.73	0.91
15 - 18	0.353	1.39	0.73	0.91

a. Based on predictive analysis of Figure 11: α = helical, β = β sheet, C = random-coil.

b. Based on probability profile of Figure 12.

c. $\langle P_\alpha \rangle$, $\langle P_\beta \rangle$, $\langle P_t \rangle$ are respectively the average conformational potential for the computed region to be in the helical, β-sheet and β-turn conformation. $\langle P_\alpha \rangle$ is the sum of the P_α values of the individual residues divided by the total residues in the segment under consideration. $\langle P_\beta \rangle$ and $\langle P_t \rangle$ are derived in a similar manner by averaging the P_β and P_t values respectively.

C-peptide structure. This result is in agreement with the speculation of Yu et al. (1972) that the connecting peptide in porcine insulin may have considerable α-helical fractions in their Raman studies of porcine insulin and proinsulin. As discussed above, the conformation of the C-peptides seems to be same as that existing in proinsulin. In other words, the C-peptide does not undergo conformational change when it is released. From Figure 12 and Table 1, it can be seen that the possibility of β structure is very rare. No sign of antiparallel-β structure is seen in the Raman spectrum. This is not in agreement with one of the manifestations made by Markussen and Schiff (1973). They proposed two possible conformations of bovine C-peptide. One of them is the prolate ellipsoid and the other, oblate type. The prolate type which accompanies an assumption of antiparallel-β structure should be eliminated. But their prediction of bending about middle position is reasonable when the predicted β turns (Table 1) are examined. Concerning the oblate type, it will be discussed in a later part of this section.

The relative probability (p_t) that a tetrapeptide will form a β turn is ploted against the amino acid sequence in Figure 12. In the case of bovine C-peptide, two regions, 15-18, and 17-20, have a β-turn probability greater than 0.5×10^{-4} and one region, 16-19, has three residues having an average frequency of occurrence of β turn greater than 0.07 among tetrapeptide residues although the p_t value is

(a)

	1	2	3	4	5	6	7	8	9	10	11	12	13
	Glu	Val	Glu	Gly	Pro	Gln	Val	Gly	Ala	Leu	Glu	Leu	Ala
α	H	h	H	B	B	[h	h	B	H	H	H	H	H]
β	B	H	B	i	b	h	H	i	I	h	B	h	I

	14	15	16	17	18	19	20	21	22	23	24	25	26
	Gly	Gly	Pro	Gly	Ala	Gly	Gly	Leu	Glu	Gly	Pro	Pro	Gln
α	B	B	B	B	H	B	B	H	H	B	B	B	h
β	i	i	b	i	I	i	i	h	B	i	b	b	h

(b)

	1	2	3	4	5	6	7	8	9	10	11	12	13
	Glu	Ala	Glu	Asn	Pro	Gln	Ala	Gly	Ala	Val	Glu	Leu	Gly
α	H	H	H	b	B	[h	H	B	H	h	H	H]	B
β	B	I	B	b	b	h	I	i	I	H	B	h	i

	14	15	16	17	18	19	20	21	22	23	24	25	26
	Gly	Gly	Leu	Gly	Gly	Leu	Gln	Ala	Leu	Ala	Leu	Glu	Gly
α	B	B	H	B	B	[H	h	H	H	H	H	H]	B
β	i	i	h	i	i	h	h	I	h	I	h	B	i

Figure 12. Predictive Analysis of Helical and Random-Coil Regions in (a) Bovine and (b) Porcine C-peptides

	27 Pro	28 Pro	29 Gln
α	B	B	h
β	b	b	h

Figure 12. (Continued)

below the cut-off value 0.5×10^{-4}. But following the rule (Chou and Fasman 1974) stating that when $p_t > 0.5 \times 10^{-4}$ for tetrapeptides at both sites j and j+1, only the site with the higher p_t value is predicted as a β turn, the region of 16-19 is not necessarily included since it has a lower p_t value (0.256×10^{-4}) than those of 15-18 (1.296×10^{-4}) and 17-20 (0.910×10^{-4}). The predicted results can be formulated as a hairpin type conformation, not forming hydrogen bonds between the segments antiparallel each other. The schematic diagram is drawn in Figure 14(a). The predicted β-turns of porcine C-peptide are figured out in Table 1. There are two tetrapeptide region having a p_t value greater than 0.5×10^{-4}; 12-15 (0.611×10^{-4}) and 13-16 (0.754×10^{-4}). By the rule mentioned above, one having lower value than the other is excluded. The region, 13-16, is considered as β turn forming tetrapeptide. The regions, 14-17, 15-18 and 17-20 do not have the p_t value higher than the cut-off line in Figure 12, but each of them has three with the average frequency greater than 0.07 out of four residues. The region of 14-17 (site j) has a lower p_t value when compared to either that of 13-16 (j-1) or 15-18 (j+1), and the region, 17-20, has three residues already allocated to α-helix region. Therefore, the regions, 12-15 and 17-20, are not included independently in the predicted β turn. The higher values of $\langle P_t \rangle$ in bovine or porcine C-peptides are an indication of the bending structure. A

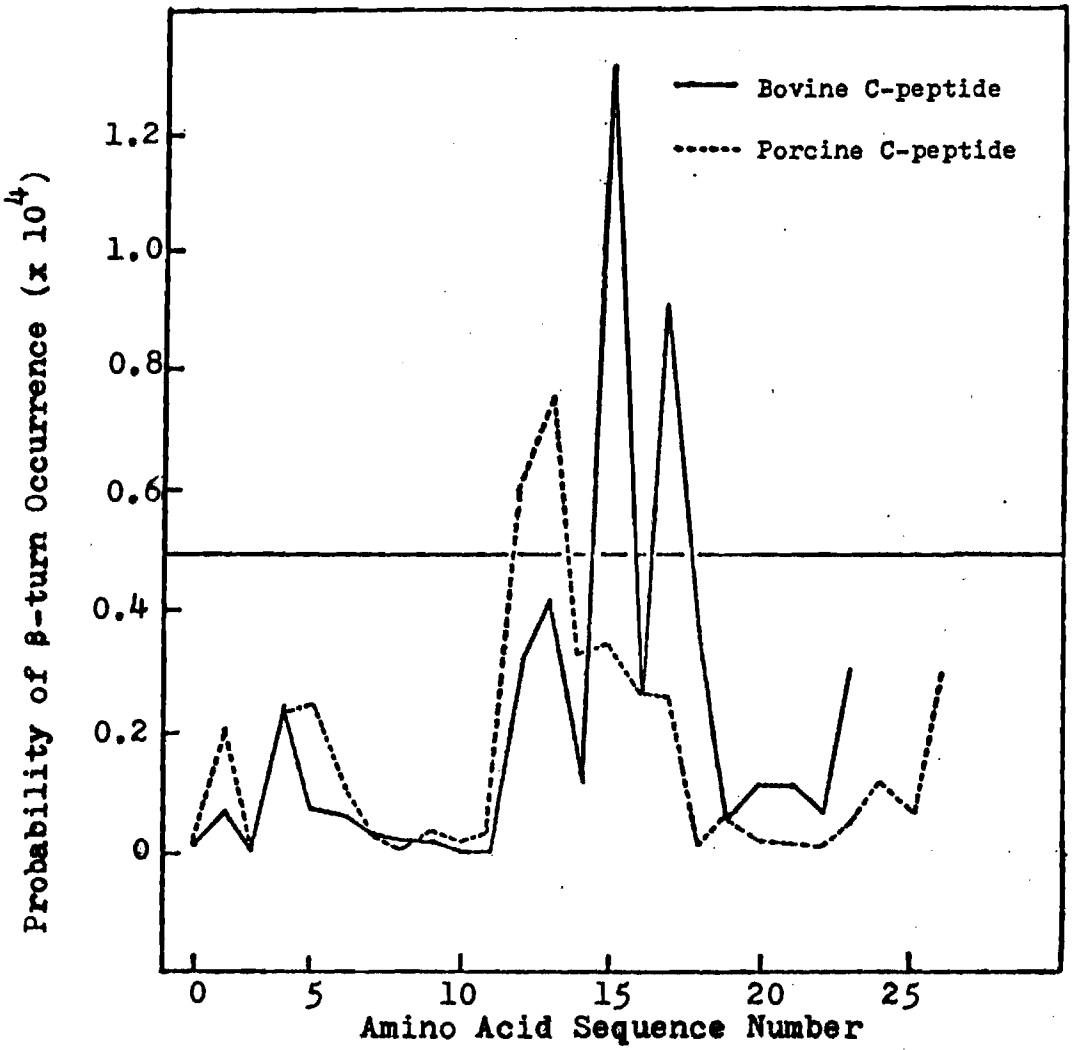

Figure 13. Probability of Tetrapeptide β-turn in C-peptides

schematic diagram of a secondary structure of porcine C-peptide is shown in Figure 14(b). The total length of bovine C-peptide can be estimated approximately using the predicted model. Since a stretched chain has a length of about 3.5 Å per amino acid residue and α helix has a length of 1.5 Å per amino acid residue (Markussen and Shiff 1973), eight amino acid residues in α helix and 17 in random-coil give a length of 75 Å. Markussen and Shiff (1973) proposed an oblate ellipsoid of revolution for bovine C-peptide as an alternative to the prolate. In that case, the diameter was estimated 33 Å. If the terminals are around 10 Å apart as shown by X-ray data (Blundell et al. 1972a, b) and two released dipeptides are regarded as having a length of 14 Å, the maximum length of bovine C-peptide becomes 99 Å. This value will agree with the oblate with a 33 Å diameter. Therefore, it is concluded that the bovine C-peptide has a conformation similar to a planar ring.

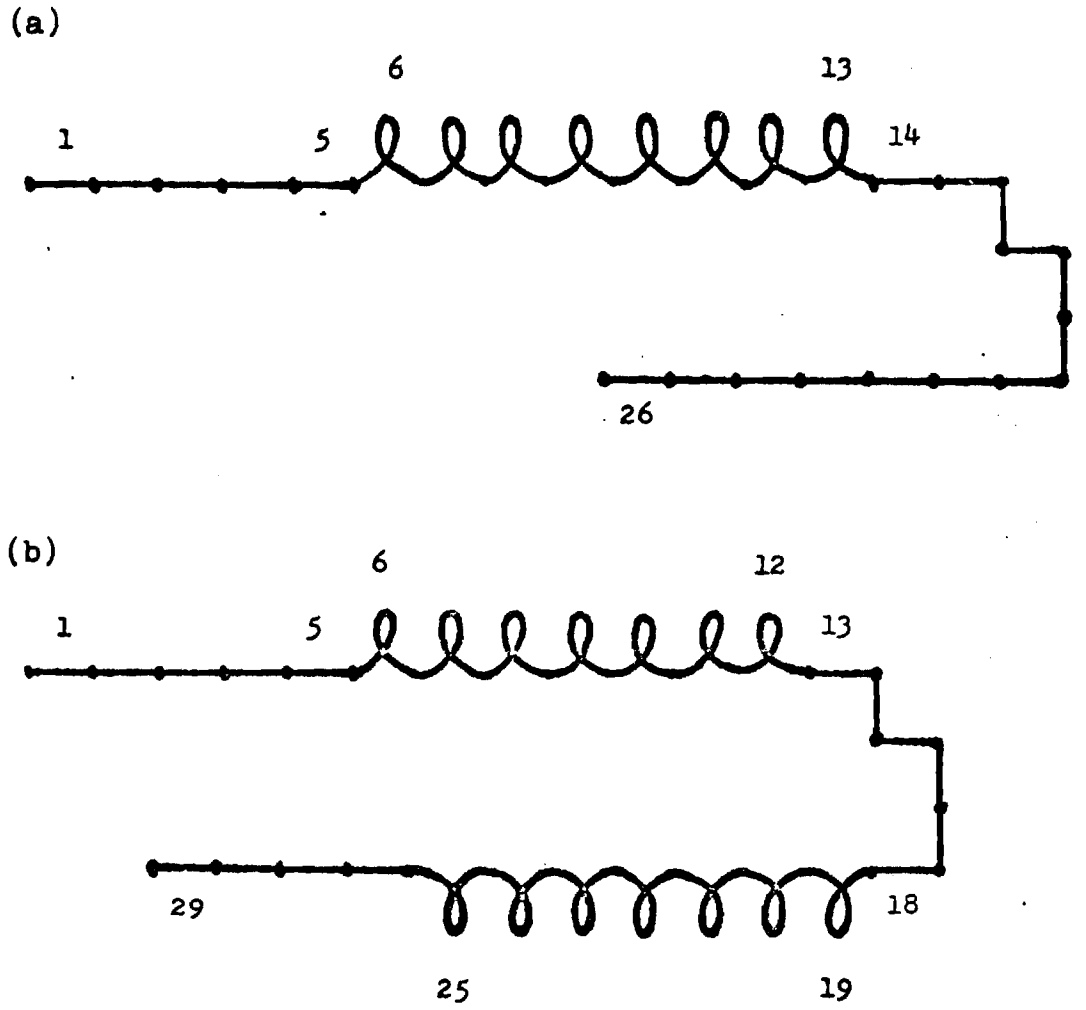

Figure 14. Schematic Diagram of Secondary Structures Predicted in (a) Bovine and (b) Porcine C-peptides

CHAPTER V

NATIVE SINGLE-CRYSTAL LYSOZYME AND LYSOZYME-UREA INTERACTION

Introduction

Urea is well known as a strong chemical denaturant. Many proteins become fully unfolded in the presence of 8 M urea. Quite unusually, it was recently reported that lysozyme could be crystallized even in 9 M urea medium. A preliminary X-ray diffraction study by Snape et al. (1974) indicated that the structure of lysozyme underwent major conformational changes up to 3 M urea and that only minor additional changes took place at higher concentration of urea. They agreed with Edelhoch and Steiner (1962) in that the molecule retained some structure that was not a random-coil even in 9 M urea.

Laser Raman spectroscopy is a useful tool for indication of various secondary structures of a protein in crystals and in solution (Yu and Jo 1973a, Chapter I of this dissertation). Here is reported the Raman spectra of lysozyme single crystals grown in various concentrations of urea under the same conditions as those used for X-ray study (Snape et al. 1974). The extent of structural alterations by urea will be discussed.

Materials and Methods

Native tetragonal crystals of hen egg white lysozyme chloride were grown following the method of Alderton and Fevold (refer to the cited references of Snape *et al.* 1974): 200 mg of lyophilized lysozyme were dissolved with thorough stirring in 2.5 ml of 0.04 M acetate buffer at pH 4.7. Then sodium chloride solution (2.5 ml, 10% W/V) was added dropwise over several minutes, the final solution was centrifuged and the supernatant were left undisturbed in $4^{o}C$ cold room for up to two weeks. Several crystals with size of about 1 mm were selected for Raman spectra.

Cocrystals of lysozyme and urea were grown according to the method described by Snape *et al.* (1974). Lysozyme (150 mg) was dissolved in 1 ml of 0.4 M acetate buffer at pH 4.7 of the urea concentrations, 6, 9 and 13.5 M. After the system reached equilibrium (1 hr.), 0.5 ml of saturated sodium chloride solution was added to each solution giving 4, 6 and 9 M urea concentrations, respectively. After further gentle stirring, the solutions were centrifuged, and the supernatant of each was left undisturbed in a $4^{o}C$ cold room for up to two weeks. It was observed that the crystal size became smaller as the concentration of urea was increased.

Hen egg white lysozyme used for sample preparation was obtained from Sigma Chemical Co. in the lyophilized state (Lot No. 72C-8060, 3 x crystallized, dialyzed and

lyophilized), sodium chloride (certified A.C.S. grade), glacial acetic acid (reagent A.C.S. grade) and sodium acetate (certified A.C.S. grade) from Fisher Scientific Co. and urea (ultra pure grade) from Schwartz/Mann. All the chemicals were used without further purification.

The Raman technique used in the handling of a single-crystal protein was described in detail previously (Yu et al. 1974, Yu and Jo 1973b). Al the spectra presented here were obtained with the 514.5 nm line of an argon laser (Coherent Radiation 52G) and at 4 cm^{-1} resolution. The Raman system and optical arrangements used were the same as those described in other works published from our laboratory (Yu 1974, Yu and East 1975).

Results and Discussion

In Figures 1(a)-(d), Raman spectra of single-crystal lysozyme grown in 0, 4, 6 and 9 M urea are presented.

The shape of the amide III band in the single-crystal spectrum (Figure 1(b)), compared to the solution or crystalline powder spectrum (Yu and Jo 1973a), is sharper. But neither any frequency shifts nor changes in relative intensities of the respective peaks are detected in the amide III region. This finding may imply that the backbone conformation of lysozyme in the single-crystal state is actually the same as that of solution or crystalline powder except that the single-crystal is in a more ordered state

since the sharpening of the amide III region has been regarded as an uniformity in the hydrogen bonds. Assignment of amide III in the lysozyme spectrum was made by Yu and Jo (1973a): 1272, 1258 and 1238 cm^{-1} to α-helices, random-coil and antiparallel-β sheet respectively. Through the spectra of Figures 1(a)-(c), no marked changes in shape and frequency number are observed in the amide III and C-C-N vibration (900-1150 cm^{-1}) regions. However, the spectrum of a single-crystal grown in 9 M urea shows a definite change in the shape of the amide III band. The present Raman data indicate that no changes in backbone conformation of lysozyme occur at concentrations up to 6 M urea.

When the spectrum of the single-crystal lysozyme in Figure 1(b) is compared to that of solution or crystal powder (Yu and Jo 1973a, Chapter I of this dissertation), the intensity ratio of the lines, 696/721, is greater in the single-crystal case. Furthermore, the ratio has increased for the single-crystals as the concentration of urea was increased to 4 and 6 M as shown in Figures 1(a) and (c). At the concentration of 9 M urea, these lines are interfered with by the bands of mother liquid. Also it is observed that the spectral line at 659 cm^{-1} has increased in intensity up to 6 M urea. The 659, 696 and 721 cm^{-1} lines were assinged to the C-S stretching vibrations of half-cystine groups in the P_H, P_N and P_C forms respectively and possibly the 696 cm^{-1} line also to the C-S stretching vibrations of two methionine

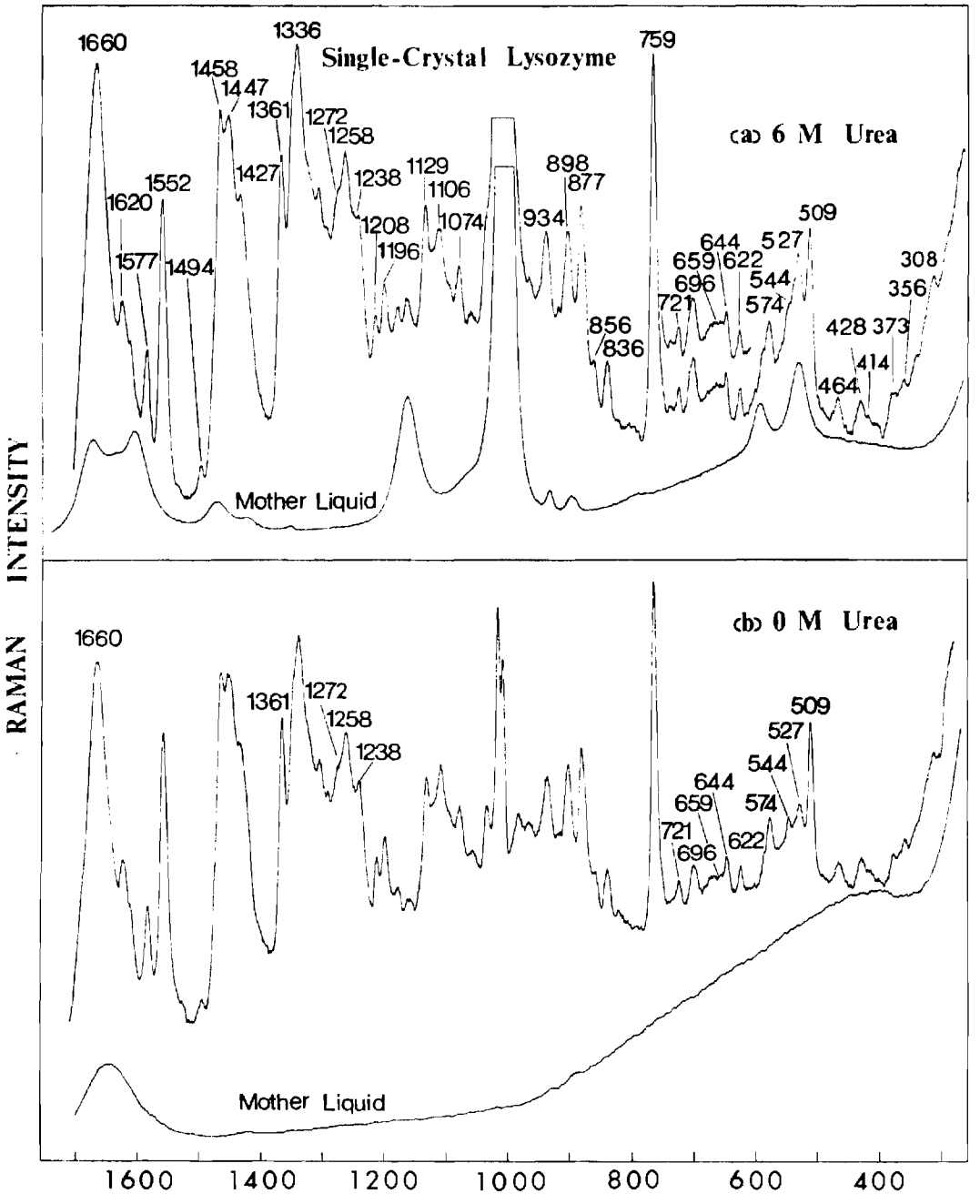

Figure 1. Raman Spectra of Single-Crystal Lysozyme in Various Concentration of Urea

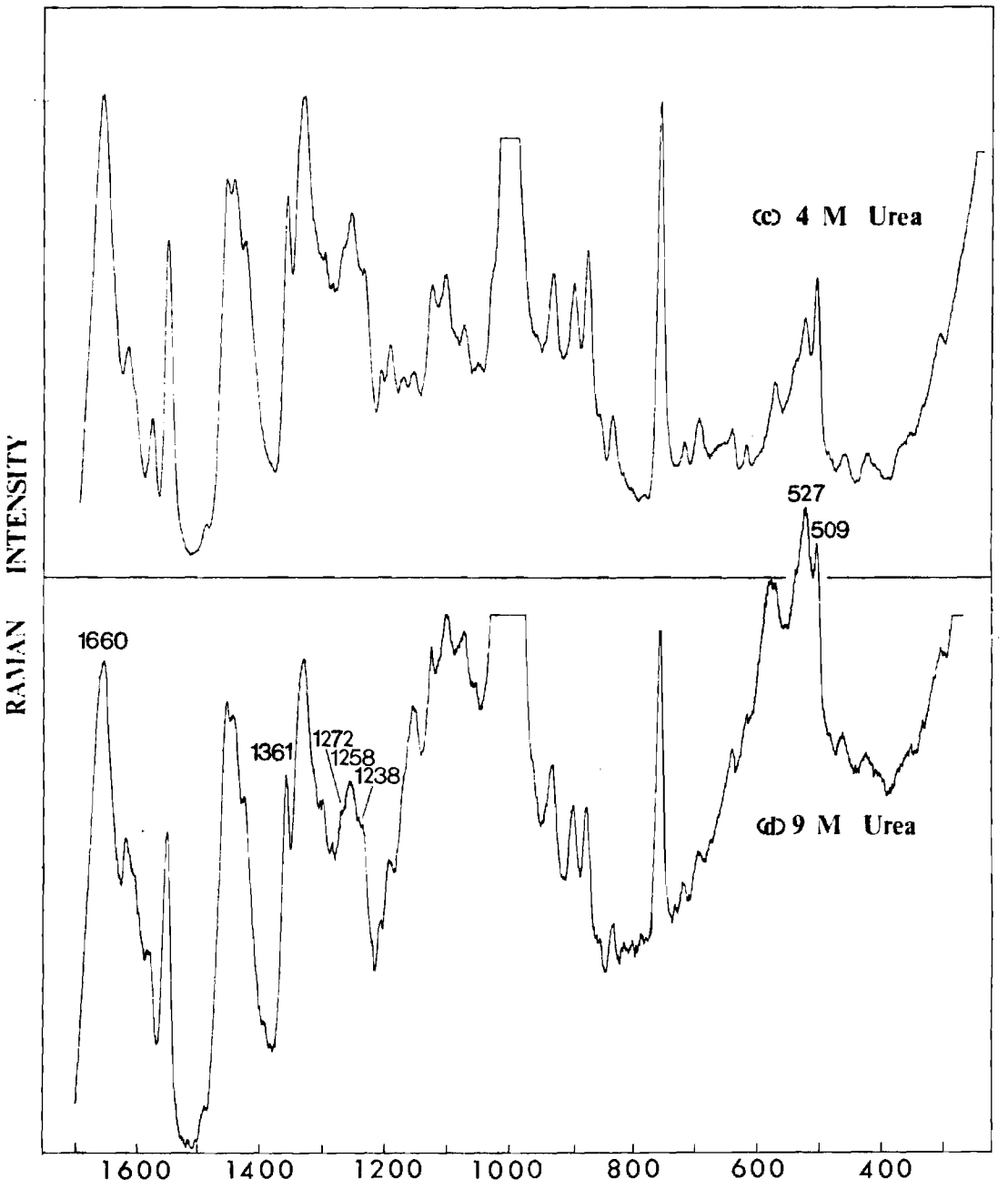

Figure 1. (Continued)

groups in the trans-guch forms (Nogami et al. 1975, Miyazawa and Sugeta 1974, Sugeta et al. 1972). Although the exact nature of these intensity variations is not known, it is quite possible to assume changes in either geometries or local environment of half-cystine groups. However, it is unlikely that the half-cystine undergo the geometrical transformation between P_H, P_C, and P_N forms without affecting the main-backbone conformation or disulfide stretching lines at 509 and 527 cm^{-1} lines arising from the three disulfide bonds in the gauch-gauch-gauch form and one in gauch-gauch-trans respectively (Sugeta et al. 1972, 1973, Miyazawa and Sugeta 1974, Nakanishi et al. 1974). Therefore, the environmental changes of the disulfide bonds, possibly the half-cystine groups in the P_N form, in urea may be caused by the conformational changes which have been detected by X-ray diffraction (Snape et al. 1974) and the measurements of other physical properties (Hamaguchi and Rokkaku 1960, Leónis 1956).

In early studies, Frank-Conrat (1951) found that the disulfide bonds of the lysozyme molecule are more readily reduced by mercaptoethanol in 8 M urea than in the absence of urea. Lapanje and Rupley (1973) measured enthalpies of reduction of disulfide bonds of lysozyme by dithiothreitol at 30°C in water and 9 M urea and found that the value in 9 M urea was less by about 3 Kcal/mole of disulfide bond than that in water. These facts may be the results that

urea has made the disulfide bonds more accessible to solvent in the highly concentrated urea solution. This phenomenon may be due to the backbone conformational changes of the protein in 9 M urea as revealed in Figure 1(d) or/and the local environmental changes of the cystine linkages.

On the other hand, the chemical activity of lysozyme has been known to be very sensitive to the concentrations even lower than 6 M urea (James and Hilborn 1958, Berthou and Jollès 1973, Hamaguchi and Rokkaku 1960, Léonis 1956). Therefore, it is assumed that the conformational changes of some side-chains are critical in chemical activity of lysozyme in the urea solution. Also, an assumption made by Léonis (1956) that lytic activity of the enzyme in urea may be due to changes in the substrate itself has not been disproved so far.

CHAPTER VI

CONCLUSIONS

The conclusions reached in various chapters can be summarized:

1. The lysozyme molecule was found to have the same backbone conformation in aqueous solution as in the crystalline state. The same is true for ribonuclease A(RNase A). However, some of the side-chains in both proteins have subtle differences in conformation between the two phases.

2. The lyophilization process alters the protein conformation of both lysozyme and RNase A.

3. It was confirmed that trichloroacetic acid(TCA) exerts a considerable destablizing effect on the conformation of proteins, such as lysozyme, RNase A and insulin, and it was suggested that denaturing state is in predominantly antiparallel-β backbone conformation.

4. The relationship between the Raman lines of the tyrosyl ring and their structural state in the protein molecule was established from model compounds and was applied to other protein systems. It was found that tyrosines in the snake venom toxins studied are involved in certain intramolecular interaction, but upon heat denaturation, they are affected.

5. The first application of Raman technique to a protein containing β-parallel component, carboxypeptidase A (CPDase A), was made with incomplete understanding. Further investigation on model systems is desirable.

6. The relationship between the Raman frequency and geometry about disulfide bond was also confirmed in insulin and CPDase A cases and was applied to snake venom proteins. It was found that the geometries about the three disulfide bonds of cobramine B are in gauch-gauch-gauch, in neurotoxin α three of four are in the gauch-gauch-gauch form and one in gauch-gauch-trans and in P. platurus major toxin and L. semifasciata toxin b all four are in the gauch-gauch-gauch form. Variation of Raman lines of the S-S stretching vibration was interpreted based on the change of geometry. No change in the integral intensity of the S-S Raman band has been observed unless the S-S bond is broken.

7. The Raman bands of the C-S stretching vibration of the protein spectra were discussed. Although the discussion is not satisfactory to the proteins which contain both half-cystine and methionine residues, it was suggested that the snake venom proteins investigated in the present research have their half-cystine groups predominantly in the P_H form. It was found that Raman intensity of the C-S stretching vibration was not only due to the geometry with regard to the C-S bond but also due to the changes in the local environment.

8. Chou and Fasmans' (1974) predictive method of protein conformation was used for the prediction of the structure of a snake venom protein, neurotoxin α. The present Raman data with the help of the conformation calculations predicted that most of snake venom proteins studied here have a considerable amount of random-coil backbone relative to α or β structures.

9. By comparing a single-crystal insulin to the air-dried powder for Raman studies, the dehydration effect caused by laser beam was manifested on the spectrum.

10. The utility of the true contour of the amide III region of a protein spectrum was suggested for a more reliable interpretation of backbone conformation.

11. Raman data of bovine C-peptide disclosed that its backbone conformation consists of α-helix and random-coil. This result is consistent with the predicted conformation. The prediction was also calculated for porcine C-peptide. They were compared to X-ray and CD results of the connecting peptide in proinsulin. It was concluded that the C-peptide may not undergo a conformational change when it is released from its precursor. No sign of β structure was noticed in either Raman data or the predicted conformation.

12. It was concluded that carboxymethylated insulin chains and insulin and glucagon fibrils have their backbone mainly in antiparallel-β structure. The interpretation of

the amide I was supported by a theoretical basis of Krimm and Abe (1972).

13. Lysozyme single-crystals were grown in high concentrations of urea. It was found that the effect of urea on lysozyme backbone conformation was not noticed up to the concentration of 6 M urea except for subtle changes in local environment of the C-S bond of the cystine linkages.

From these results, it can be seen that laser Raman spectroscopy of proteins has produced interesting and significantly new structural information. Yet much of the work requires amplification and refinement.

In this work, studies have been focused on non-resonance Raman scattering of protein molecules. In principle, resonance Raman spectra of proteins can be obtained when one uses UV laser (200-400 nm) irradiation. The selective resonance enhancement ($\sim 10^3$) of Raman scattering is observed when the exciting laser frequency is tuned to the absorption region of peptide group, Phe, Trp, Tyr, S-S, S-H or His.

The reason for employing the Raman spectroscopy for the conformational analysis of proteins, when X-ray diffraction, optical rotatory dispersion (ORD) and circular dichroism (CD) are available, is that Raman measurements can determine the protein structure with equal ease in a variety of phases such as in solid, gel and solution whereas X-ray diffraction structure determinations are restricted to

crystals and ORD and CD can be obtained only in solution. Serious discrepancies between estimates of main conformational components, i.e., α-helix, β-sheet and random-coil contents by CD compared to X-ray have been reported for several proteins (Barela and Darall 1974, Chou and Fasman 1974, Chen et al. 1972). Laser Raman technique can be used to determine whether these differences are caused by deficiencies in interpretations of the CD or are due to conformational changes accompanying dissolution. Moreover, the information obtained by the Raman technique is more specific and selective in protein structure determination than other spectroscopic techniques are. The present conclusions are worthwhile to biochemists who try to establish structure-function relationships. They often rely on the crystal structure of proteins determined by the X-ray method to explain the mechanisms of biochemical reactions which mostly occur in water media. The difficulty in the growing of protein crystals and the interpretation of the data has restricted the X-ray technique to a very limited number of proteins.

At the present stage, the Raman works on proteins are still qualitative and procedure for quantitative estimation of various secondary structure of a protein remains to be worked out. The future development of laser Raman sepctroscopy appears to be the studies of drug-protein interaction, inhibitor-protein complex, chromosomal structure, membrane structure and intact biological specimens.

APPENDIX

Spectroscopic Factors on Line Intensity and Shape*

The scattering intensity of Raman transition from the initial state n to the final state n is given by

$$I_{mn} = \frac{2^3\pi}{3^3c^4} I_o (\nu_o - \nu_{mn})^4 \sum_{ij} (\alpha_{ij})^2_{mn} \qquad (1)$$

Here I_o and ν_o are the intensity and frequency of the exciting light, ν_{mn} is the vibrational frequency and $(\alpha_{ij})_{mn}$, ijth component of scattering tensor α.

Connection between the scatteing tensor and the quantum theory of the interaction between matter and radiation is effected by the Kramers-Heisenbers-Dirac dispersion equation

$$(\alpha_{ij})_{mn} = \frac{1}{h} \sum_e \left[\frac{(M_j)_{me}(M_i)_{en}}{\nu_e - \nu_o + i\delta_e} - \frac{(M_i)_{me}(M_j)_{en}}{\nu_e + \nu_o + i\delta_e} \right] \qquad (2)$$

The notation $(M_i)_{en}$ means the transition moment from the ground state n to a given excited state e along the direction i. The same way of description is applied to $(M_j)_{me}$, $(M_i)_{me}$ and $(M_j)_{en}$. δ_e is a damping constant. When $\nu_e >> \nu_o$

---

* Ref. (Spiro 1974, Banwell 1972).

for all excited states, equation (2) is independent of exciting frequency for all excited states. This is the principle of non-resonance Raman. When ν_o approaches ν_e, the first term on the right-hand side of equation (2) is to dominate the scattering tensor giving the preresonance enhancement. In the case that $\nu_o - \nu_e$ is very small, which means that the exciting line falls on electronic absorption region, the first term of equation (2) is greatly enhanced giving resonance Raman effect of certain Raman mode if the transition moments are sizable. In the present study, the Raman effect has been monitored in the non-resonance region at a fixed exciting frequency. Therefore, the dependence of scattering intensity on the exciting frequency was not considered. In contributing to the variation of Raman intensity, there are several other factors which are inherent to the molecular state. They include concentration, local environment, intra and intermolecular interaction and hyper- and hypochromism of Raman active groups. Raman hypochromism were proposed and interpreted (Pézolet et al. 1975, Tomlinson and Peticolas 1970). Considered in analyzing the Raman intensity of protein spectra are concentration, local environment and inter or intramolecular interaction of Raman active normal modes.

In shaping the Raman band, slit width as an instrumental factor and natural line width and degree of uniformness in bond or geometry of the Raman active group as molecular

factors are considered. Although the slit width is vital to the resolving power of the spectrometer, it is not counted in the present work since the experimental setting of the slit width was small enough to detect the natural line width of the Raman active modes. The natural line width is contributed by collision broadening, Doppler broadening and Heisenberg uncertainty principle. In the region of frequency shift (10^{12}-10^{14}HZ), the uncertainty ($\delta\nu = 10^{8}$HZ) which are calculated under the assumption that the scattering process occurs within 10^{-8} sec. is small enough to be neglected. It is hard to expect that the changes in sample phase makes noticeable differences in the natural linewidth arising from the collision broadening and Doppler broadening for macromolecules. Furthermore, the collision broadening contribution is regarded as negligibly small for biopolymers relative to the case of small molecules. Therefore, only considered here as an important factor is the degree of uniformness in geometry and bond formation of the Raman active group. The examples revealed in the protein study are amide I, amide III and S-S stretching vibrations. The nature of the characteristic group frequencies of amide I and amide III has been studied theoretically in N-methylacetamide by Miyazawa et al. (1958). On the basis of their results they assigned the amide I and III frequencies, respectively, to C=O stretching and one of the coupled vibrations of C-N stretching and N-H in-plane bending. The

other coupled vibration appears near 1567 cm^{-1} (amide III); it is extremely weak in the Raman effect but quite strong in infrared absorption. The uniformness in the hydrogen-bond formation between the N-H and C=O groups is known to play a role in band width of the amide bands.

LITERATURE CITED*

Alderton, G. and Fevold, J., J. Biol. Chem. 164, 1 (1946).

Ambrose, J. A. and Laskowski, M. Science 115, 358 (1952).

Banwell, C. N. in Fundamentals of Molecular Spectroscopy, pp. 19-25, McGraw-Hill, London (1972).

Barela, T. D. and Darnall, D. W., Biochemistry 13, 1694 (1974).

Berthous, J. and Jollès, P., Fed. Eur. Biochem. Soc. Lett. 31, 189 (1973).

Blake, C. D. F., Mair, G. A., North, A. C. T., Phillips, D. C. and Sarma, V. R., Proc. Roy. Soc. (London), Ser. B 167, 365 (1967).

Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A. and Vijayan, M., Nature 231, 506 (1971).

Blundell, T. L., Cutfield, J. F., Dodson, E. J., Dodson, G. G., Hodgkin, D. C. and Mercola, D. A., in Cold Spring Harbor Symp. Quant. Biol. 36, 233 (1972).

Blundell, T. L., Dodson, G., Hodgkin, D. and Mercola, D., Advan. Prot. Chem. 26, 279 (1972).

Blout, E. R., Deloze, C. and Asadourian, A., J. Am. Chem. Soc. 83, 1895 (1961).

Borsook, H., U. S. Pat. 2510543, June 6 (1950).

Botes, D. P. and Strydom, D. J., J. Biol. Chem. 244, 4147 (1969).

Braganca, B. M., Patel, N. T. and Badrinath, P. G., Biochim. Biophys. Acta 136, 508 (1967).

---

* Journal title abbreviations used are listed in "List of Periodicals", Chemical Abstracts, 1961.

Brunner, H. and Sussner, H., Biochim. Biophys. Acta 271, 16 (1972).

Burke, M. J. and Rougvie, M. A., Biochemistry 11, 2435 (1972).

Chance, R. E., Ellis, R. M. and Bromer, W. W., Science 161, 165 (1968).

Chang, C. C., Yang, C. C., Hamaguchi, K., Nakai, K. and Hayashi, K., Biochim. Biophys. Acta 236, 164 (1971).

Chen, M. C., Lord, R. C. and Mendelsohn, R., Biochim. Biophys. Acta 328, 252 (1973).

Chen, M. C., Lord, R. C. and Mendelsohn, R., J. Am. Chem. Soc. 96, 3038 (1974).

Chen, Y. H., Yang, J. T. and Martinez, H. M., Biochemistry 11, 4120 (1972).

Chicheportiche, R., Rochat, C., Sampieri, F. and Lazdunski, M., Biochemistry 11, 1681 (1972).

Chou, P. Y. and Fasman, G. D., Biochemistry 13, 222 (1974).

Clark, J. L., Cho, S., Rubenstein, A. H. and Steiner, D. F., Biochem. Biophys. Res. Comm. 35, 456 (1969).

Dickerson, R. E. and Geis, I., in The Structure and Action of Proteins p. 87, Harper and Row, New York (1969).

Edelhoch, H. and Steiner, R. F., Biochim. Biophys. Acta. 60, 365 (1962).

Edsall, J. T., in Structural Chemistry and Molecular Biology (Rich, A., and Davidson, N., eds) p. 88, Freeman, San Francisco (1968).

Fanconi, B., Tomlinson, B., Nafie, L. A., Small, E. W. and Peticolas, W. L., J. Chem. Phys. 51, 3993 (1969).

Frank, B. H. and Veros, A. J., Biochem. Biophys. Res. Comm. 32, 155 (1969).

Frank-Conrat, H., J. Am. Chem. Soc. 73, 625 (1951).

Fullerton, W. W., Potter, R. and Low, B. W., Proc. Nat. Acad. Sci. U. S. 66, 1213 (1970).

Grant, P. T., Coombs, T. L. and Frank, B. H., Biochem. J. 126, 433 (1972).

Hamaguchi, K. and Rokkaku, K., J. Biochem. 48, 358 (1960).

Hamaguchi, K. and Kurono, A., J. Biochem. 54, 111 (1963).

Hayashi, K., Takechi, M. and Sasaki, T., Biochem. Biophys. Res. Comm. 45, 1357 (1971).

Hnilica, L. S., John, E. W. and Butler, J. A. V., Biochem. J. 82, 123 (1962).

Hong, B. S. and Tu, T. T., Fed. Proc. 29, 888 (1970).

James, L. K. and Hilborn, J. A., Biochem. Biophys. Acta 151, 279 (1968).

Johansen, J. T. and Vallee, B. L., Proc. Nat. Acad. Sci. U. S. 68, 2532 (1971).

John, E. W. and Butler, J. A. V., Biochem. J. 82, 15 (1962).

Kartha, G., Bello, J. and Harker, D., Nature 213, 862 (1967).

Kemmler, W., Peterson, J. D. and Steiner, D. F., J. Biol. Chem. 246, 6786 (1971).

Ko, A., Smyth, D. G., Markussen, J. and Sundby, F., Europ. J. Biochem. 20, 190 (1971).

Koenig, J. L., J. Polymer Sci. Part D, 59 (1972).

Krimm, S., J. Mol. Biol. 4, 528 (1962).

Krimm, S. and Abe, Y., Proc. Nat. Acad.Sci. U. S. 69, 2788 (1972).

Lapanje, S. and Rupley, J. A., Biochemistry 12, 2370 (1973).

Larsen, P. R. and Wolff, J., J. Biol. Chem. 243, 1282 (1968).

Lèonis, J., Arch. Biochem. Biophys. 65, 182 (1956).

Lewin, S., Abst. Fed. Eur. Biochem. Soc. 6th Meeting, Madrid, Spain, p. 1065 (1969).

Lewin, S., J. Theor. Biol. 26, 481 (1970).

Lewin, S. and Rosmarin, M. N., Biochem. J. 119, 629 (1970).

Lipscomb, W. N., Hartsuck, J. A., Reeke, G. N., Quiocho, F. A., Bethge, P. H., Ludwig, M. L., Steitz, T. A., Muirhead, H. and Coppola, J. C., Brookhaven Symp. Biol. 21, 24 (1968).

Lord, R. C., Pure Appl. Chem. Suppl. 7, 179 (1971).

Lord, R. C. and Mendelsohn, R., J. Am. Chem. Soc. 94, 2133 (1972).

Lord, R. C. and Yu, N. T., J. Mol. Biol. 50, 509 (1970a).

Lord, R. C. and Yu, N. T., J. Mol. Biol. 51, 203 (1970b).

Markussen, J. Int. J. Protein Res. 3, 201 (1971).

Markussen, J. and Schiff, H. E., Int. J. Peptide Protein Res. 5, 69 (1973).

McDonald, M. R., Methods Enzymol. 2, 427 (1955).

Miyazawa, T., J. Chem. Phys. 32, 1647 (1960).

Miyazawa, T., Shimanouchi, T. and Mizushima, S., J. Chem. Phys. 29, 611 (1958).

Miyazawa, T. and Sugeta, H., Abst. US-Japan Joint Raman Seminar, Cleveland, Ohio (1974).

Nakanishi, M., Tsuboi, M. and Ikegami, A., J. Mol. Biol. 70, 351 (1972).

Nogami, N., Sugeta, H. and Miyazawa, T., Chem. Lett. 147 (1975).

Oyer, P. E., Cho, S. and Steiner, D. F., Fed. Proc. 29, 533 (1970).

Perutz, M. F., J. Mol. Biol. 14, 646 (1965).

Petra, P. H., Methods Enzymol. 14, 460 (1970).

Pezolet, M., Yu, T-J and Peticolas, W. L., J. Raman Spectrosc. 3, 55 (1975).

Phillips, D. C., Sci. Amer. 215, 78 (1966).

Pimentel, G. C. and McClellan, A. L., The Hydrogen Bond, p. 293, Freeman, San Francisco (1960).

Quiocho, F. A. and Lipscomb, W. N., Advan. Prot. Chem. 25, 1 (1971).

Raymond, M. L. and Tu, A. T., Biochim. Biophys. Acta, 285, 498 (1972).

Richards, R. E. and Thompson, H. W., J. Chem. Soc., p. 1248 (1947).

Rougvie, M. A. and Shriver, C. N., Biophysical Society Abst. 17th Annual Meeting, Feb. 27-Mar. 2, 1973, p. 204a (1973).

Rupley, J. A., in Structure and Stability of Biological Macromolecules (Timasheff, S. N. and Fasman, G.D., eds) pp. 291-352, Dekker, New York (1969).

Salokangas, A., Smyth, D. G., Markussen, J. and Sundby, F., Europ. J. Biochem. 20, 183 (1971).

Schlichtkrull, J., Acta Chem. Scand. 10, 1459 (1956).

Schulz, G. G., Barry, C. D., Friedman, J. Chou, P. Y., Fasman, G. D., Finkelstein, A. V., Lim, V. I., Ptitsyn, O. B., Kabat, E. A., Wu. T. T., Levitt, M., Robson, B. and Nakano, K., Nature 250, 140 (1974).

Seto, A., Sato, S. and Tamiya, N., Biochim. Biophys. Acta 214, 483 (1970).

Shugar, D., Biochem. J. 52, 142 (1952).

Small, E. W., Fanconi, B. and Peticolars, W. L., J. Chem. Phys. 52, 4369 (1970).

Snape, K. W., Tjian, R., Blake, C. C. F. and Koshland, Jr., D. E., Nature 250, 295 (1974).

Spiro, T. G., in Chemical and Biological Applications of Lasers, Vol. I, p. 29, Ed. Moor, C. B., Acad. Press, N. Y. (1974).

Steiner, D. F., Cho, S., Oyer, P. E., Terris, S., Peterson, J. D. and Rubenstein, A. H., J. Biol. Chem. 246, 1365 (1971).

Steiner, D. F. and Clark, J. L., Proc. Nat. Acad. Sci. U. S. 60, 622 (1968).

Sugeta, H., Go, A. and Miyazawa, T., Chem. Lett. 83 (1972).

Sugeta, H., Go, A. and Miyazawa, T., Bull. Chem. Soc. Japan 46, 3407 (1973).

Takenaka, A., Takenaka, O., Mizota, T., Shibata, K., Inada, Y., J. Biochem. 70, 63 (1971).

Tanford, C. and Hauenstein, J. D., J. Am. Chem. Soc. 78, 1287 (1956).

Tanford, C. and Hauenstein, J. D., and Rands, D. G., J. Am. Chem. Soc. 77, 6409 (1955).

Tomlinson, B. and Peticolas, W. L., J. Chem. Phys. 52, 2154 (1970).

Tu, A. T., Ann. Rev. Biochem. 42, 235 (1973).

Tu, A. T., J. Agr. Food Chem. 22, 36 (1974).

Tu, A. T. and Hong, B., J. Biol. Chem. 246, 2772 (1971).

Tu, A. T., Hong, B. and Solie, T. N., Biochemistry 10, 1295 (1971).

Tu, A. T., Lin, T. A., Jo, B. H., Yu, N. T. and Bieber, A. L., Biochemistry, submitted (1975).

Wallach, D. F. H., Graham, J. M. and Oseroff, A. R., Fed. Eur. Biochem. Soc. Lett. 7, 330 (1970).

Wauch, D. F., J. Am. Chem. Soc. 66, 663 (1944).

Woody, R. W., Friedman, M. E. and Scheraga, H. A., Biochemistry 5, 2034 (1966).

Wyckoff, H. W., Hardman, K. D., Allewell, N. M., Inagami, T., Johnson, L. N. and Richards, F. M., J. Biol. Chem. 242, 3984 (1967).

Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B. and Richards, F. M., J. Biol. Chem. 245, 305 (1970).

Yang, C. C., Toxicon 3, 19 (1965).

Yang, C. D., Biochim. Biophys. Acta 133, 346 (1967).

Yu, N. T., Ph.D. Thesis, Massachusetts Institutes of Technology (1969).

Yu, N. T., J. Am. Chem. Soc. 96, 4664 (1974).

Yu, N. T. and East, E. J., J. Biol. Chem. 250, 2196 (1975).

Yu, N. T. and Jo, B. H., Arch. Biochem. Biophys. 156, 469 (1973a).

Yu, N. T. and Jo, B. H., J. Am. Chem. Soc. 95, 5033 (1973b).

Yu, N. T., Jo, B. H., Chang, R. C. C. and Huber, J. D., Arch. Biochem. Biophys. 160, 614 (1974).

Yu, N. T., Jo, B. H. and Liu, C. S., J. Am. Chem. Soc. 94, 7572 (1972).

Yu, N. T., Jo, B. H. and O'Shea, D. C., Arch. Biochem. Biophys. 156, 71 (1973).

Yu, N. T., Lin, T. S. and Tu, A. T., J. Biol. Chem. 250, 1782 (1975).

Yu, N. T. and Liu, C. S., J. Am. Chem. Soc. 94, 5127 (1972).

Yu, N. T., Liu, C. S. and O'Shea, D. C., J. Mol. Biol. 70, 117 (1972).

VITA

Byeong Hyeok Jo was born on November 16, 1944, Seoul, Korea, the son of Young-Soon and In-Hwa Jo. He attended Seoul High School and received his B.S. in 1967 and M.S. in 1970 from Chungang University, Seoul, Korea. During undergraduate studies, he was awarded the University President Scholarship and the Yupoong Fellowship. From 1967 to 1971, he worked with the Korea Tungsten Mining Company. He entered Georgia Institute of Technology in June 1971, in order to attain his Doctor of Philosophy in the School of Chemistry.

In March 1972, he married Soon-Ae Kim and they are parents of a daughter, Ellen.

His research was done under the direction of Dr. Nai-Teng Yu. During the course of study, he was supported by DeWitt Wallace Fellowship and a research assistantship.